

143428

From: Chan, Christina  
Sent: Wednesday, January 26, 2005 2:45 PM  
To: Ramirez, Delia; STIC-Biotech/ChemLib  
Subject: RE: rush search 09/886400

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644  
(571)-272-0841  
Remsen, 3E89

CRFE

-----Original Message-----

From: Ramirez, Delia  
Sent: Wednesday, January 26, 2005 2:39 PM  
To: Chan, Christina  
Subject: rush search 09/886400

Hi Christina,

Please approve the following interference search: SEQ ID NO: 4 in the protein and nucleic acid databases

Thank you very much,

Delia M. Ramirez, Ph.D.  
Patent Examiner  
Recombinant Enzymes-Art Unit 1652  
USPTO  
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70  
Alexandria, VA 22314  
(571) 272-0938  
delia.ramirez@uspto.gov

\*\*\*\*\*

STAFF USE ONLY

Searcher: [Signature]  
Searcher Phone: 2-2504  
Date Searcher Picked up: 1/27/05  
Date Completed: 2/1/05  
Searcher Prep/Rev. Time: 10  
Online Time: 15

\*\*\*\*\*

Type of Search

NA Sequence: # ✓  
AA Sequence: # ✓  
Structure: # ✓  
Bibliographic:             
Litigation:             
Patent Family:             
Other:           

\*\*\*\*\*

Vendors and cost where applicable

STN:             
DIALOG:             
QUESTEL/ORBIT:             
LEXIS/NEXIS:             
SEQUENCE SYSTEM: ✓  
WWW/Internet:             
Other(Specify):

This Page Blank (uspto)



```

QY 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
DB 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
QY 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
DB 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

## RESULT 2

```

US-09-886-400-4
; Sequence 4, Application US/09886400
; Patent No. US20020045226A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-09-886-400-4

```

## Query Match

```

Best Local Similarity 99.5%; Score 1870; DB 9; Length 364;
Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRALVFHGNLYQYAEIPKSEIPKVIKRAYIPVETLKEIPGINTGTTLKFLPKDIIID 60
DB 1 LRALVFHGNLYQYAEIPKSEIPKVIKRAYIPVETLKEIPGINTGTTLKFLPKDIIID 60
QY 61 LVKGGIASDLIELIGTSYTHALPLPLSRVEAQVQREKVEKELEFSPKGFWLPBELAY 120
DB 61 LVKGGIASDLIELIGTSYTHALPLPLSRVEAQVQREKVEKELEFSPKGFWLPBELAY 120
QY 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
DB 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
QY 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
DB 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360

```

```

DB 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

## RESULT 3

```

US-10-112-357-4
; Sequence 4, Application US/10112357
; Publication No. US20020115099A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,357
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-357-4

```

```

Query Match
Best Local Similarity 99.5%; Score 1870; DB 13; Length 364;
Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRALVFHGNLYQYAEIPKSEIPKVIKRAYIPVETLKEIPGINTGTTLKFLPKDIIID 60
DB 1 LRALVFHGNLYQYAEIPKSEIPKVIKRAYIPVETLKEIPGINTGTTLKFLPKDIIID 60
QY 61 LVKGGIASDLIELIGTSYTHALPLPLSRVEAQVQREKVEKELEFSPKGFWLPBELAY 120
DB 61 LVKGGIASDLIELIGTSYTHALPLPLSRVEAQVQREKVEKELEFSPKGFWLPBELAY 120
QY 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
DB 121 DPTIIPALIKONGYEVYLFADGEAMLFSAHLSAIPKIPKLYPHLIKAQEKRFRTSYLLG 180
QY 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTILKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
DB 301 DKSRLIWEDEGNARLNLMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

## RESULT 4

```

US-10-114-403-4
; Sequence 4, Application US/10114403
; Publication No. US20020115100A1
; GENERAL INFORMATION:

```

APPLICANT: DIVERSA CORPORATION  
APPLICANT: Murphy, Dennis  
APPLICANT: Ried, John  
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: DIVER1120-4  
CURRENT APPLICATION NUMBER: US/10/114,403  
PRIORITY FILING DATE: 2002-04-01  
PRIORITY FILING DATE: 2001-06-20  
PRIORITY FILING DATE: 2000-07-19  
PRIORITY FILING DATE: 2000-07-19  
PRIORITY FILING DATE: 2000-07-19  
PRIORITY FILING DATE: 1999-09-20  
PRIORITY FILING DATE: 1996-03-08  
PRIORITY FILING DATE: 1996-03-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 364  
TYPE: PR  
ORGANISM: Thermococcus alcaliphilus  
US-10-114-403-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
Best Local Similarity 99.5%; Pred. No. 5,7e-165;  
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINTGYTLKFLPKDIID 60  
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINTGYTLKFLPKDIID 60  
QY 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREVEKEBELFELSPKGFWLPETAY 120  
DB 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREVEKEBELFELSPKGFWLPETAY 120  
QY 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
DB 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
QY 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
DB 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
QY 181 LRELKRAIKLVFEGSKVTLLKAVKOIEAVPVVAVTAVMLGIGRLPLNNPKKVASWIEDKD 240  
DB 181 LRELKRAIKLVFEGSKVTLLKAVKOIEAVPVVAVTAVMLGIGRLPLNNPKKVASWIEDKD 240  
QY 241 NILLYGTDIEFIGYRDLAGYMSVEGLLEVIDELNSLCLPSELKHSGRELYLRTSSWAP 300  
DB 241 NILLYGTDIEFIGYRDLAGYMSVEGLLEVIDELNSLCLPSELKHSGRELYLRTSSWAP 300  
QY 301 DKSIRIRREDEGNARLNLMLSYNMGELAFLENSDARGMEPLPERRLDAPRAIYNDWGE 360  
DB 301 DKSIRIRREDEGNARLNLMLSYNMGELAFLENSDARGMEPLPERRLDAPRAIYNDWGE 360  
QY 361 NGEP 364  
DB 361 NGEP 364

RESULT 5  
US-10-116-606-4  
Sequence 4, Application US/10116606  
Publication No. US20020119515A1  
GENERAL INFORMATION:  
APPLICANT: DIVERSA CORPORATION  
APPLICANT: Murphy, Dennis  
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH  
FILE REFERENCE: DIVER1120-4  
CURRENT APPLICATION NUMBER: US/10/116,606  
PRIORITY FILING DATE: 2002-04-03  
PRIORITY FILING DATE: US/09/886,400  
PRIORITY FILING DATE: 2001-06-20  
PRIORITY FILING DATE: 1996-03-08  
PRIORITY FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/407,806  
PRIORITY FILING DATE: 1999-09-20  
PRIORITY FILING DATE: 1999-09-20  
PRIORITY FILING DATE: 1996-03-08  
PRIORITY FILING DATE: 1996-03-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 364  
TYPE: PR  
ORGANISM: Thermococcus alcaliphilus  
US-10-116-606-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
Best Local Similarity 99.5%; Pred. No. 5,7e-165;  
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINTGYTLKFLPKDIID 60  
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINTGYTLKFLPKDIID 60  
QY 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREVEKEBELFELSPKGFWLPETAY 120  
DB 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREVEKEBELFELSPKGFWLPETAY 120  
QY 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
DB 121 DPTIPALIKONGEYELFADGEAMLSAHLNSAIKPIKLYPHLILKAOREKFRYISYLLG 180  
QY 181 LRELKRAIKLVFEGSKVTLLKAVKOIEAVPVVAVTAVMLGIGRLPLNNPKKVASWIEDKD 240  
DB 181 LRELKRAIKLVFEGSKVTLLKAVKOIEAVPVVAVTAVMLGIGRLPLNNPKKVASWIEDKD 240  
QY 241 NILLYGTDIEFIGYRDLAGYMSVEGLLEVIDELNSLCLPSELKHSGRELYLRTSSWAP 300  
DB 241 NILLYGTDIEFIGYRDLAGYMSVEGLLEVIDELNSLCLPSELKHSGRELYLRTSSWAP 300  
QY 301 DKSIRIRREDEGNARLNLMLSYNMGELAFLENSDARGMEPLPERRLDAPRAIYNDWGE 360  
DB 301 DKSIRIRREDEGNARLNLMLSYNMGELAFLENSDARGMEPLPERRLDAPRAIYNDWGE 360  
QY 361 NGEP 364  
DB 361 NGEP 364

RESULT 6  
US-10-112-331-4  
Sequence 4, Application US/10112331  
Publication No. US20020119550A1  
GENERAL INFORMATION:  
APPLICANT: DIVERSA CORPORATION  
APPLICANT: Murphy, Dennis  
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE 1  
FILE REFERENCE: DIVER1120-4  
CURRENT APPLICATION NUMBER: US/10/112,331  
PRIORITY FILING DATE: 2002-03-29  
PRIORITY FILING DATE: 2001-06-20  
PRIORITY FILING DATE: 2001-06-20  
PRIORITY FILING DATE: 2000-07-19  
PRIORITY FILING DATE: 2000-07-19  
PRIORITY FILING DATE: 1999-09-20  
PRIORITY FILING DATE: 1996-03-08  
PRIORITY FILING DATE: 1996-03-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 364  
TYPE: PR  
ORGANISM: Thermococcus alcaliphilus  
US-10-112-331-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
Best Local Similarity 99.5%; Pred. No. 5,7e-165;  
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
DB 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIMEDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFAIYNDWGE 360
DB 301 DKSIRIMEDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFAIYNDWGE 360
QY 361 NGEF 364
DB 361 NGEF 364
```

## RESULT 7

```
US-10-112-377-4
; Sequence 4, Application US/10112377
; Publication No. US20020120108A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-377-4
```

Query Match 99.5%; Score 1870; DB 13; Length 364;  
Best Local Similarity 99.5%; Pred. No. 5,7e-165;  
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
DB 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
```

```
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIMEDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFAIYNDWGE 360
DB 301 DKSIRIMEDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFAIYNDWGE 360
QY 361 NGEF 364
DB 361 NGEF 364
```

## RESULT 8

```
US-10-116-581-4
; Sequence 4, Application US/10116581
; Publication No. US20020137116A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVER1120-4
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-116-581-4
```

Query Match 99.5%; Score 1870; DB 13; Length 364;  
Best Local Similarity 99.5%; Pred. No. 5,7e-165;  
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
DB 1 LRALVHFGNLOVAEIPKSEIPKVIKAYIPVITLLIKEEIPFGNITGTYLKLPDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQVORDREVKELFEVSPKGFMLPELAY 120
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILIKQREKRFYISYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEFIgyrdIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
```

QY 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 |||||  
 Db 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 QY 361 NGEP 364  
 |||||  
 Db 361 NGEP 364

## RESULT 9

US-10-112-442-4  
 ; Sequence 4, Application US/10112442  
 ; Publication No. US20020150997A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,442  
 ; PRIOR FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: 09/886,400  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR APPLICATION NUMBER: 08/613,220  
 ; PRIOR FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-442-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99.5%; Pred. No. 5.7e-165;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFGNLOAYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPKDID 60  
 |||||  
 Db 1 LRALVFGNLOAYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPKDID 60  
 QY 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFWLPDELAY 120  
 |||||  
 Db 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFWLPDELAY 120  
 QY 121 DPPIPALKONGEYELPADGEAMLFSAHLSAIPKIPLYPHLIKAKREKFRYISYLLG 180  
 |||||  
 Db 121 DPPIPALKONGEYELPADGEAMLFSAHLSAIPKIPLYPHLIKAKREKFRYISYLLG 180  
 QY 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLNMPKKVASWIEDKD 240  
 |||||  
 Db 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLNMPKKVASWIEDKD 240  
 QY 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300  
 |||||  
 Db 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300  
 QY 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 |||||  
 Db 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 QY 361 NGEP 364  
 |||||  
 Db 361 NGEP 364

RESULT 10  
 US-10-112-418-4

; Sequence 4, Application US/10112418  
 ; Publication No. US20020155486A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,418  
 ; PRIOR FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: 09/886,400  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR APPLICATION NUMBER: 08/613,220  
 ; PRIOR FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-418-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99.5%; Pred. No. 5.7e-165;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFGNLOAYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPKDID 60  
 |||||  
 Db 1 LRALVFGNLOAYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPKDID 60  
 QY 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFWLPDELAY 120  
 |||||  
 Db 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFWLPDELAY 120  
 QY 121 DPPIPALKONGEYELPADGEAMLFSAHLSAIPKIPLYPHLIKAKREKFRYISYLLG 180  
 |||||  
 Db 121 DPPIPALKONGEYELPADGEAMLFSAHLSAIPKIPLYPHLIKAKREKFRYISYLLG 180  
 QY 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLNMPKKVASWIEDKD 240  
 |||||  
 Db 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLNMPKKVASWIEDKD 240  
 QY 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300  
 |||||  
 Db 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300  
 QY 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 |||||  
 Db 301 DKSIRIRREDEGNARLMLSYNNRGEFLAENSDARGWEPLPERRLDAPRAIYNDWGE 360  
 QY 361 NGEP 364  
 |||||  
 Db 361 NGEP 364

## RESULT 11

US-10-114-083-4  
 ; Sequence 4, Application US/10114083  
 ; Publication No. US20020160464A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/114,083  
 ; CURRENT FILING DATE: 2002-04-01

```

; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-114-083-4

```

```

Query Match          99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAKYIPVIEITLKEEIPFGINTGYTLKPLKPID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAKYIPVIEITLKEEIPFGINTGYTLKPLKPID 60
QY 61 LVKGIASDLIEITIGTSTYTHAIPPLPLSRVEAOVQDRVEKEELFEVSPKGFPLBLAY 120
DB 61 LVKGIASDLIEITIGTSTYTHAIPPLPLSRVEAOVQDRVEKEELFEVSPKGFPLBLAY 120
QY 121 DPPIPAIKDNGEYELPADGEAMLFSALHNSAIPKIPPLVPHLIKQREKFRYISYLLG 180
DB 121 DPPIPAIKDNGEYELPADGEAMLFSALHNSAIPKIPPLVPHLIKQREKFRYISYLLG 180
QY 121 DPPIPAIKDNGEYELPADGEAMLFSALHNSAIPKIPPLVPHLIKQREKFRYISYLLG 180
DB 121 DPPIPAIKDNGEYELPADGEAMLFSALHNSAIPKIPPLVPHLIKQREKFRYISYLLG 180
QY 181 LRELKAIKLVESGKVTLKAVQDIEAVPVVAVTAVMLGIGRLPLNPKKVASWIDKD 240
DB 181 LRELKAIKLVESGKVTLKAVQDIEAVPVVAVTAVMLGIGRLPLNPKKVASWIDKD 240
QY 181 LRELKAIKLVESGKVTLKAVQDIEAVPVVAVTAVMLGIGRLPLNPKKVASWIDKD 240
DB 181 LRELKAIKLVESGKVTLKAVQDIEAVPVVAVTAVMLGIGRLPLNPKKVASWIDKD 240
QY 241 NILLYGDIIEFIRGDIAGYMSVEGLEEVDELINSELCPESELKHSGRELYLTSSWAP 300
DB 241 NILLYGDIIEFIRGDIAGYMSVEGLEEVDELINSELCPESELKHSGRELYLTSSWAP 300
QY 301 DKSILIRFEDGNAFLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
DB 301 DKSILIRFEDGNAFLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

# RESULT 12

```

US-10-282-122A-53281
; Sequence 53281, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyt, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA, 0344
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23,578
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53281
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-282-122A-53281

```

```

Query Match          5.9%; Score 111.5; DB 15; Length 890;
Best Local Similarity 23.6%; Pred. No. 0.57;
Matches 72; Conservative 48; Mismatches 98; Indels 87; Gaps 16;

```

```

QY 53 FLPKDIIIDLVKGIASDLIEITIGTSTYTHAIPPLPLSRVEAOVQDRVEKEELFEVSPK 112
DB 562 FQP-DLIIAVGGSSANDACKIMWVMEHPEVFDOLA-----MFMDIRKRVY-VFPR- 612
QY 113 FWLPELAYPIIP-----AII--KNGEYELPADGEAMLFSALHNSAIPKIP 158
DB 613 --MGEKAYPAALPTSAGTSEVTPFAVITDQSGVKYPLADYELMNMALIIDADMMEP 670
QY 159 LYPHL-----IKQREKFRYISYL-----LGLRELKAIKLVEE-----GKVTL 198
DB 671 --PRLTAAGVDALTHALEAYVSMLETERPADGAL-----QAGKIIEYLPRAVKGNKNDK 724
QY 199 KAVQDIEAVPVVAVTAVMLGIGRLPLNPKKVASWIDKD-----P 229
DB 725 EAREKMAASTWAGMFAFNAFLGICHSIAHKLGAFFHVGAVANALLINEVIFKNCAP 784
QY 230 KYVASWIEDK--DNILLYGDIIEFIRGDIAGY--RMSVEGLEEVDELINSELCPESEL 285
DB 785 NKMGAFSQYRYPDCIORYA---EFASFAGIKSGTDOEKVDNLKALIDELKAKVGLPKTIK 841
QY 286 HSGRE 290
DB 842 EAGVE 846

```

# RESULT 13

```

US-10-369-493-23237
; Sequence 23237, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```



SEQ ID NO 23237  
LENGTH: 312  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
US-10-369-493-23237

Query Match 5.8%; Score 109; DB 14; Length 312;  
Best Local Similarity 23.2%; Pred. No. 0.22;

Matches 44; Conservative 32; Mismatches 56; Indels 56; Gaps 9;

QY 10 LQVAFIPSEIPK-----VIEKAYIPVETLKEIPGKLNITGYTLKFLPKDIDL 61  
DB 130 MTVAAYKSGFPKERVIGSGVLTDR--PRTFAER--LNLG-----VKDVGTF 175  
QY 62 VKGASDLIRIGSYTHAI--LPLPLSRVEAOVQDREYKEELFEVSPKG--FWLPE 117  
DB 176 VLGGGDDMVPLVRYSYAGIPLFTLIPKRIDALVETRRKGGGIVMLGNGSAVYAPA 235  
QY 118 LAYDPIIPALIKDN-----GYEYLF-----ADGEAMLFS-----A 147  
DB 236 ASLTEMVAAILKQDRRULPTAYLGEYGYEGITVGVTVGNGLEQIIELELTDERA 295  
QY 148 HNSAIKPIK 157  
DB 296 QLNKSVESVK 305

RESULT 14  
US-10-425-115-291380

Sequence 291380, Application US/10425115  
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425.115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 291380

LENGTH: 590

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(590)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577\_28825C.1.pap

US-10-425-115-291380

Query Match 5.6%; Score 106; DB 17; Length 590;  
Best Local Similarity 23.3%; Pred. No. 1;

Matches 59; Conservative 46; Mismatches 94; Indels 54; Gaps 13;

QY 22 KVEKAYIPVETLIK-----EIPFGNITGYTLKFLPKDIDLKGGIASDLIEITG 76  
DB 343 KKAGGTYHVALDKFTKMEYKRIASLTSAKAVEFI--QDI--FRGINSITIDGS 399  
QY 77 SYTHAIL-----DLPLSRVEAOVQ-----DREVEELFEVSPK--GFWLP 116  
DB 400 NPTSESEFDEORSIQIKVASLHPRANGQVERANGMIPEALRKRVDPKKEKFGKWIR 459  
QY 117 ELADPIIPALIKONGEYLF--ADGEA-----MFSAHNSAIKPIKFLPHLIXA 166  
DB 460 ELPTVVVLPADLRGAPRLVESTAEATRELYIDVEERLNTVQSAR--YQOTLR 517  
QY 167 QREKFRYISYLLGLRELKRAIKLVFEGSKVTLKAVKDIKAVPVW--VAVNTAVMLGIGR 223  
DB 518 YHDKTIRHRSFVQDVLVRI--LTGBGRHKL-----PQWBGPRMVSEVTRPGSYR 567

QY 224 LPLNPKKVA--SW 235  
DB 568 LTMQDTEVGNNSW 580

RESULT 15

US-10-424-599-175517

Sequence 175517, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424.599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 175517

LENGTH: 573

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12950C.1.pap

US-10-424-599-175517

Query Match 5.6%; Score 105.5; DB 15; Length 573;  
Best Local Similarity 21.3%; Pred. No. 1.1;

Matches 77; Conservative 57; Mismatches 97; Indels 131; Gaps 19;

QY 6 FHGNL-QYAEIPKS-----EIPKVIKAYIPV-----IETL 35  
DB 290 FLGSLGKLGVPSSLISGSEVPPIVSPFSPYCYCPCPGISTCPSIAVNGSPNSIETL 349  
QY 36 IKKEIPFGNITGYTLKFLPKDIDLKGGIASDLIRIGSYTHAILPLPLSRVEAOV 95  
DB 350 ---PFPGASLNLPLS-----VNLL-----DPVPLGTSMDFP--PFLP--DPLV 388  
QY 96 QDRREVEBELFEVSPKGFMLPELAYDPI--IPAI--LMDNGEYEFADGEAMLFSALNSA 152  
DB 389 RMSLPTSOQIPTFTP-----LMCDPIVHVPIVDVCSGGGYLVSHGPAVSPS----- 435  
QY 153 IKPIKPLYPHLIKAQREKFRYISYLLGLRELKRAIKLVFEG-----KVTLKAVKD 203  
DB 436 ---IPPLHPNLVKRLIPESDAVV-----KGARETLRLLSGSSQGNQMMRDITLPAI-- 484  
QY 204 IEAVPVVAVNTAVMLGIGRLPLMNPVKVASMIEDKDNILLYGTDFIEFIGYDI----- 257  
DB 485 -----LTNPD-----ENQNNIIVAGSRGLYGTGRDINALANS 516  
QY 258 --AGRMSVSGLEVIDELNSELCLPSELKHSREGLVLRSSMAPDLSLRIMRDEGNAR 315  
DB 517 IAAAGLVSLSGVSKVDSGVYSEL-----ENYGNLXAVKNSN--DSGGGAFPLEDEGSS 568  
QY 316 LN 317  
DB 569 LD 570

Search completed: January 27, 2005, 05:59:41  
Job time : 65 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 05:58:26 ; Search time 22 Seconds  
(without alignments)  
1097.261 Million cell updates/sec

Title: US-09-886-400A-4  
Perfect score: 1879  
Sequence: 1 LRALVFHGNQYAEIPKSEI.....RLDAPRAIYNDMGNGE 364

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	100.0	364	US-10-166-606-4	Sequence 4, Appli
2	1879	100.0	364	US-10-112-211A-4	Sequence 4, Appli
3	1589	84.6	346	US-08-613-230B-4	Sequence 4, Appli
4	140.5	7.5	647	US-07-894-212A-8	Sequence 8, Appli
5	140.5	7.5	649	US-07-894-212A-2	Sequence 2, Appli
6	140.5	7.5	650	US-07-893-928A-1	Sequence 1, Appli
7	105	5.6	227	US-09-107-532A-5554	Sequence 5554, Ap
8	101.5	5.4	360	US-09-710-279-2150	Sequence 2150, Ap
9	101.5	5.4	376	US-09-710-279-2106	Sequence 2106, Ap
10	101	5.4	329	US-08-270-013B-2	Sequence 2, Appli
11	101	5.4	329	US-08-838-418-2	Sequence 2, Appli
12	97	5.2	653	US-09-540-226-3128	Sequence 3128, Ap
13	94.5	5.0	367	US-09-134-001C-3678	Sequence 3678, Ap
14	92	4.9	375	US-09-710-279-1840	Sequence 1840, Ap
15	92	4.9	845	US-09-248-796A-15394	Sequence 15394, A
16	91.5	4.9	741	US-09-252-991A-31448	Sequence 31448, A
17	91.5	4.9	1009	US-09-693-146-4	Sequence 4, Appli
18	91.5	4.8	1031	US-09-543-681A-8245	Sequence 8245, Ap
19	90.5	4.8	406	US-09-134-001C-3570	Sequence 3570, Ap
20	90.5	4.8	789	US-09-248-796A-19294	Sequence 19294, A
21	90.5	4.8	1171	US-09-248-796A-16043	Sequence 16043, A
22	89.5	4.8	322	US-09-488-039A-12383	Sequence 12383, A
23	89	4.7	857	US-09-248-796A-20522	Sequence 20522, A
24	88.5	4.7	399	US-09-252-991A-23741	Sequence 23741, A
25	88.5	4.7	764	US-09-235-451-36	Sequence 36, Appli
26	88.5	4.7	764	US-09-978-303-36	Sequence 36, Appli
27	88.5	4.7	993	US-09-894-998A-50	Sequence 50, Appli

28	88.5	4.7	1037	US-09-894-998A-54	Sequence 54, Appli
29	88.5	4.7	1113	US-09-894-998A-51	Sequence 51, Appli
30	88.5	4.7	2089	US-08-418-893D-23	Sequence 23, Appli
31	88.5	4.7	2089	US-08-418-893D-24	Sequence 24, Appli
32	88	4.7	329	US-09-710-279-1460	Sequence 1460, Ap
33	88	4.7	339	US-09-328-352-1144	Sequence 1144, Ap
34	88	4.7	339	US-09-489-039A-11436	Sequence 11436, A
35	88	4.7	420	US-09-248-796A-23483	Sequence 23483, A
36	88	4.7	1151	US-09-710-279-2448	Sequence 2448, Ap
37	88	4.7	1154	US-09-134-001C-3428	Sequence 3428, Ap
38	87.5	4.7	689	US-08-248-021A-2	Sequence 2, Appli
39	87	4.6	235	US-09-141-135-2	Sequence 2, Appli
40	87	4.6	235	US-09-533-029-98	Sequence 98, Appli
41	87	4.6	657	US-09-355-166-1	Sequence 1, Appli
42	87	4.6	657	US-09-462-845-2	Sequence 2, Appli
43	86.5	4.6	796	US-09-540-236-2265	Sequence 2265, Ap
44	86	4.6	1068	US-09-966-997-11	Sequence 11, Appli
45	85.5	4.6	441	US-09-489-039A-9546	Sequence 9546, Ap

## ALIGNMENTS

```
RESULT 1
US-10-166-606-4
Sequence 4, Application US/10166606
Patent No. 6644756
GENERAL INFORMATION:
APPLICANT: Murphy, Dennis
APPLICANT: Reid, John
TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
FILE REFERENCE: 09010-004005
CURRENT APPLICATION NUMBER: US/10/166,606
PRIOR FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: US 09/407,806
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 06/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-166-606-4
Query Match 100.0%; Score 1879; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.9e-191;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LRALVFHGNQYAEIPKSEIPKVEKAYIVIEITLKEEIPFGINTGYTLKFLPKDIIID 60
1 LRALVFHGNQYAEIPKSEIPKVEKAYIVIEITLKEEIPFGINTGYTLKFLPKDIIID 60
Db 1 LRALVFHGNQYAEIPKSEIPKVEKAYIVIEITLKEEIPFGINTGYTLKFLPKDIIID 60
QY 61 LVKGIASDLIEITIGISTYTAIILPLPLSRVEAOVQDREYKEELFEVSPKGFMLPLAY 120
61 LVKGIASDLIEITIGISTYTAIILPLPLSRVEAOVQDREYKEELFEVSPKGFMLPLAY 120
Db 61 LVKGIASDLIEITIGISTYTAIILPLPLSRVEAOVQDREYKEELFEVSPKGFMLPLAY 120
QY 121 DPTIIPALKNQGEYVLPADDEAMLFSAHLSAIPKIPKLPVHLIKAKREKFRITYSLIG 180
121 DPTIIPALKNQGEYVLPADDEAMLFSAHLSAIPKIPKLPVHLIKAKREKFRITYSLIG 180
Db 121 DPTIIPALKNQGEYVLPADDEAMLFSAHLSAIPKIPKLPVHLIKAKREKFRITYSLIG 180
QY 181 LRELKRAIKLVFEGSKVTLKAVKDIKIAVAVVAVNTAVMLGIGRLPLNNPKVSWIBDKD 240
181 LRELKRAIKLVFEGSKVTLKAVKDIKIAVAVVAVNTAVMLGIGRLPLNNPKVSWIBDKD 240
Db 181 LRELKRAIKLVFEGSKVTLKAVKDIKIAVAVVAVNTAVMLGIGRLPLNNPKVSWIBDKD 240
QY 241 NILLYGDIIEFIRGDIAGYRMSVEGLLEVIDENSLCLPSBKHSGRGLYLTSTSWAP 300
241 NILLYGDIIEFIRGDIAGYRMSVEGLLEVIDENSLCLPSBKHSGRGLYLTSTSWAP 300
Db 241 NILLYGDIIEFIRGDIAGYRMSVEGLLEVIDENSLCLPSBKHSGRGLYLTSTSWAP 300
QY 301 DKSLIRREDEGNARLMLSYNNRGEFLAENSNDAGWPELPRRLDAFRAIYNDMGE 360
301 DKSLIRREDEGNARLMLSYNNRGEFLAENSNDAGWPELPRRLDAFRAIYNDMGE 360
```

Db 301 DKSRIWREDEGNARLNMLSYNMRGELAFLAENSDARGWEPPLERRLDAPRAIYNDWRGE 360  
QY 361 NGEF 364  
Db 361 NGEF 364

RESULT 2  
US-10-112-231A-4  
; Sequence 4, Application US/10112231A  
; Patent No. 6744246  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; APPLICANT: Reid, John  
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR  
; FILE REFERENCE: 09010-004005  
; CURRENT APPLICATION NUMBER: US/10/112,231A  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: US 09/407,806  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 08/613,220  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 4  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Thermococcus alcaliphilus  
US-10-112-231A-4

Query Match 100.0%; Score 1879; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4,9e-191;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRAVFGNLOQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINITYTLKFLPKDIID 60  
Db 1 LRAVFGNLOQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINITYTLKFLPKDIID 60  
QY 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVEAQVORDREYKEELFEVSPKGFMLPELAY 120  
Db 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVEAQVORDREYKEELFEVSPKGFMLPELAY 120  
QY 121 DPPIPAIKKONGYEYLPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180  
Db 121 DPPIPAIKKONGYEYLPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180  
QY 181 LRELRAIKLVEBEGKVTLKAVKDI EAVPVWAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240  
Db 181 LRELRAIKLVEBEGKVTLKAVKDI EAVPVWAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240  
QY 241 NILLYGTDIEFIIGRDIAGRMSVEGLLEVIDELINSELCLPSELKSGRELYLRTSSWAP 300  
Db 241 NILLYGTDIEFIIGRDIAGRMSVEGLLEVIDELINSELCLPSELKSGRELYLRTSSWAP 300  
QY 301 DKSRIWREDEGNARLNMLSYNMRGELAFLAENSDARGWEPPLERRLDAPRAIYNDWRGE 360  
Db 301 DKSRIWREDEGNARLNMLSYNMRGELAFLAENSDARGWEPPLERRLDAPRAIYNDWRGE 360  
QY 361 NGEF 364  
Db 361 NGEF 364

RESULT 3  
US-08-613-220B-4  
; Sequence 4, Application US/08613220B  
; Patent No. 5958751  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; APPLICANT: Reid, John  
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE  
; NUMBER OF SEQUENCES: 4

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson, P.C.  
;; STREET: 4225 Executive Square, Suite 1400  
;; CITY: La Jolla  
;; STATE: CA  
;; COUNTRY: US  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: Windows95  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/613,220B  
;; FILING DATE: 08-MAR-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hallie, Ph.D., Lisa A.  
;; REGISTRATION NUMBER: 38,347  
;; REFERENCE/DOCKET NUMBER: 09010/004001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 619-678-5070  
;; TELEFAX: 619-68-5099  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 346 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
US-08-613-220B-4

Query Match 84.6%; Score 1589; DB 2; Length 346;  
Best Local Similarity 95.1%; Pred. No. 3.1e-160;  
Matches 346; Conservative 0; Mismatches 0; Indels 18; Gaps 18;

QY 1 LRAVFGNLOQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINITYTLKFLPKDIID 60  
Db 1 LRAVFGNLOQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINITYTLKFLPKDIID 57  
QY 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVEAQVORDREYKEELFEVSPKGFMLPELAY 120  
Db 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVEAQVORDREYKEELFEVSPKGFMLPELAY 114  
QY 121 DPPIPAIKKONGYEYLPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180  
Db 121 DPPIPAIKKONGYEYLPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 171  
QY 181 LRELRAIKLVEBEGKVTLKAVKDI EAVPVWAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240  
Db 181 LRELRAIKLVEBEGKVTLKAVKDI EAVPVWAVNTAVMLGIGRLPLMNPCKVASWIEDKD 228  
QY 241 NILLYGTDIEFIIGRDIAGRMSVEGLLEVIDELINSELCLPSELKSGRELYLRTSSWAP 300  
Db 241 NILLYGTDIEFIIGRDIAGRMSVEGLLEVIDELINSELCLPSELKSGRELYLRTSSWAP 285  
QY 301 DKSRIWREDEGNARLNMLSYNMRGELAFLAENSDARGWEPPLERRLDAPRAIYNDWRGE 360  
Db 301 DKSRIWREDEGNARLNMLSYNMRGELAFLAENSDARGWEPPLERRLDAPRAIYNDWRGE 342  
QY 361 NGEF 364  
Db 361 NGEF 346

RESULT 4  
US-07-894-212A-8  
; Sequence 8, Application US/07894212A  
; Patent No. 536883

```

; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEFAX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-894-212A-8

Query Match      7.5%; Score 140.5; DB 1; Length 647;
Best Local Similarity 22.8%; Pred. No. 8.6e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYIVIEITLKEIP--FGLNITGYTLKFLP--KDIIIDVKGGIASDLIEIIGTSY 78
DB 27 EKCWMPLETL--EYENMKVAIHTSGPLIEWLQDNRPYIDLRLSLVKGQVEIVAGF 84
QY 79 THALPLPLSRVAQVORREVE--ELFEVSPKGFVLPBELADPIIPALIKONGEYVL 136
DB 85 YEPVLASIP--KEDRIEQIRLMKEMAKSIGFDARGVWLTERVQWOPELVTKLKSGIDYV 141
QY 137 FADGEAMLFSANLSAIPKIPLY-PHLIKARE-----KFRYISYLLGRELARKA 187
DB 142 IVD-----DYHFSAGLSKEELWYPTTEDEGVIAVFPIDELK--YIIPRPVDKV 192
QY 188 IKL--VFEKVTIKAV--KDIEAVPVVAVNTAVMLGIGRLPLMPKCVASWI----- 236
DB 193 LEYHSLIDGDESKVAVFHDGKEKFGIWPGETYEWY-----EKGWLRPEFDR 239
QY 237 ---EDKNILLYGDI-----FIGYRDIAGRMSVSGLEVIDELNSELCLP----- 281
DB 240 ISSDEKINLMYLEYLEKYKRGVLVPIASV-----FEM-----SEWSLPAKQARLF 287
QY 282 ---SELKHSG---RELYLRTSSWAPDKSLRIWREDEGN---ARLNLSTYNNKGELAF 330
DB 288 VEFNELKVGKIPFKYKRVFVRGGI--KNF-PYKYPSNYMHKMLMVSGLVANN----- 339
QY 331 AENS DARGWEP LPERRLDAFPAIYND--WRG 359
DB 340 -----PEARKYTLRAQCNDAYMHG 358

```

```

RESULT 5
US-07-894-212A-2
; Sequence 2, Application US/07894212A
; Patent No. 536683
; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEFAX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-894-212A-2

Query Match      7.5%; Score 140.5; DB 1; Length 649;
Best Local Similarity 22.8%; Pred. No. 8.7e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYIVIEITLKEIP--FGLNITGYTLKFLP--KDIIIDVKGGIASDLIEIIGTSY 78
DB 29 EKCWMPLETL--EYENMKVAIHTSGPLIEWLQDNRPYIDLRLSLVKGQVEIVAGF 86
QY 79 THALPLPLSRVAQVORREVE--ELFEVSPKGFVLPBELADPIIPALIKONGEYVL 136
DB 87 YEPVLASIP--KEDRIEQIRLMKEMAKSIGFDARGVWLTERVQWOPELVTKLKSGIDYV 143
QY 137 FADGEAMLFSANLSAIPKIPLY-PHLIKARE-----KFRYISYLLGRELARKA 187
DB 144 IVD-----DYHFSAGLSKEELWYPTTEDEGVIAVFPIDELK--YIIPRPVDKV 194
QY 188 IKL--VFEKVTIKAV--KDIEAVPVVAVNTAVMLGIGRLPLMPKCVASWI----- 236
DB 195 LEYHSLIDGDESKVAVFHDGKEKFGIWPGETYEWY-----EKGWLRPEFDR 241
QY 237 ---EDKNILLYGDI-----FIGYRDIAGRMSVSGLEVIDELNSELCLP----- 281
DB 242 ISSDEKINLMYLEYLEKYKRGVLVPIASV-----FEM-----SEWSLPAKQARLF 289
QY 282 ---SELKHSG---RELYLRTSSWAPDKSLRIWREDEGN---ARLNLSTYNNKGELAF 330
DB 290 VEFNELKVGKIPFKYKRVFVRGGI--KNF-PYKYPSNYMHKMLMVSGLVANN----- 341

```



Db 98 SVLTKD---GVINNTLNTNNHIEMSQLRAI---LRQCEPFLNKAHVILNAGQMSVLK 151  
Qy 295 TSSWAPDKSLIFMREDEGNALNTLSYMRGELAFLANSDARGEPERPLDAFRILY 354  
Db 152 KSD--EDPAUSILLVDEGQIQHVKVLSQNLTE-AMLMENLKKEGY-----ADVKKOLILY 201  
Qy 355 NDWRGENG 362  
Db 202 VEWSEKRG 209

## RESULT 8

US-09-710-279-2150  
Sequence 2150, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PUS480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2150  
LENGTH: 360  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2150

Query Match 5.4%; Score 101.5; DB 4; Length 360;  
Best Local Similarity 19.1%; Pred. No. 0.047;  
Matches 57; Conservative 52; Mismatches 117; Indels 73; Gaps 12;

Qy 4 LVFHGNQVAFIPKSELPKVIKAYIVETLKEEIPFGNTGYTLKPLPKDIIIDLVK 63  
Db 61 LIFKFG-----VKRIVEDGYGSIIRKLIONNI--NLIALHTNLDVNPKNVNRMLA 107  
Qy 64 GGASDLIEIGT--SYTHAIIPLPLSRVDAQVQRDEVEBELFEVSPKFWMLPELAYD 121  
Db 108 DQGLENIISMINTSSYYKVQTFIPKVIIE-----DFKDSINEL----- 147  
Qy 122 PIIPAILKNDGYEYLFADGEMLFSAHLNSAIKPIKPLYPHLIKAQEKRRFRYISYLLGL 181  
Db 148 ----GLAKEGNYEYCFPESSEG-----KQGFKEVGDASFYIGKLD-----IEYV--- 187  
Qy 182 RELRKAIKLVF----EGKVTLKAVKDIE--AVPV--WVAVNTAVMLGIGRLPLMPKCV 232  
Db 188 ----DEIKLEFMKIDNELETKRAIILDNHPEYTFVPDFPIKKNKSEVGLGIIGQLNQTM 243  
Qy 233 ASWIEDKDNILLYGTDIIEFIGYRDIAGRMSVEGLELV-----IDELNSELCLPSELKH 286  
Db 244 LDFSEYAKQKQNLIPSVRYTGQHDSPIKKVAIIIGSGGIFREYKASQIGADVFTGDIKH 302

## RESULT 9

US-09-710-279-2106  
Sequence 2106, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PUS480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2106  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2106

Query Match 5.4%; Score 101.5; DB 4; Length 376;  
Best Local Similarity 19.1%; Pred. No. 0.05;  
Matches 57; Conservative 52; Mismatches 117; Indels 73; Gaps 12;

Qy 4 LVFHGNQVAFIPKSELPKVIKAYIVETLKEEIPFGNTGYTLKPLPKDIIIDLVK 63  
Db 47 LIFKFG-----VKRIVEDGYGSIIRKLIONNI--NLIALHTNLDVNPKNVNRMLA 93  
Qy 64 GGASDLIEIGT--SYTHAIIPLPLSRVDAQVQRDEVEBELFEVSPKFWMLPELAYD 121  
Db 94 DQGLENIISMINTSSYYKVQTFIPKVIIE-----DFKDSINEL----- 133  
Qy 122 PIIPAILKNDGYEYLFADGEMLFSAHLNSAIKPIKPLYPHLIKAQEKRRFRYISYLLGL 181  
Db 134 ----GLAKEGNYEYCFPESSEG-----KQGFKEVGDASFYIGKLD-----IEYV--- 173  
Qy 182 RELRKAIKLVF----EGKVTLKAVKDIE--AVPV--WVAVNTAVMLGIGRLPLMPKCV 232  
Db 174 ----DEIKLEFMKIDNELETKRAIILDNHPEYTFVPDFPIKKNKSEVGLGIIGQLNQTM 229  
Qy 233 ASWIEDKDNILLYGTDIIEFIGYRDIAGRMSVEGLELV-----IDELNSELCLPSELKH 286  
Db 230 LDFSEYAKQKQNLIPSVRYTGQHDSPIKKVAIIIGSGGIFREYKASQIGADVFTGDIKH 288

## RESULT 10

US-08-270-013B-2  
Sequence 2, Application US/08270013B  
Patent No. 5686294  
GENERAL INFORMATION:  
APPLICANT: Sogabe et al.  
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Volt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 61601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/270,013B  
FILING DATE: 01-JUL-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 164701/1993  
FILING DATE: 02-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Robert F.  
REGISTRATION NUMBER: 27555  
REFERENCE/DOCKET NUMBER: 62321  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
TELEX: (25)3533  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-270-013B-2

Query Match 5.4%; Score 101; DB 1; Length 329;  
Best Local Similarity 25.3%; Pred. No. 0.046;  
Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;

QY 10 LQYAEIPKSEIPK-----VIEKAYIPVETLKEIPFGINITYTLKFLPKDII DL 61  
DB 130 MTYVFKESGFPKNRVIGSGVLDTR----FRTFVAEE-----LNIS-----VKDVTGF 175  
QY 62 VKGIASDLIEIGTSYTHAI--LPILPSRVEAQVQDRVEKEELFEVSPKG--FWLPE 117  
DB 176 VLGGHGDWVPLVNRYSVAGIPLKELIPKRLDAIVERTRKGGGEIVNLGNSAYYAPA 235  
QY 118 LAYDPIIPALIKDN-----GYEYLF 137  
DB 236 ASLVEMVEALIKQRRILPALVLEGEYEGEIV 269

## RESULT 11

US-08-838-418-2  
Sequence 2, Application US/08838418  
Patent No. 5744342  
GENERAL INFORMATION:  
APPLICANT: Sogabe et al.  
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE  
NUMBER OF INVENTION: DEHYDROGENASE ACTIVITY  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Leydig, Voit & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,418  
FILING DATE: 17-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/270,013  
FILING DATE: 01-JUL-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 164701/1993  
FILING DATE: 02-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Allen E.  
REGISTRATION NUMBER: 37354  
REFERENCE/DOCKET NUMBER: 78339  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
TELEX: (25)3533  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-418-2

Query Match 5.4%; Score 101; DB 1; Length 329;  
Best Local Similarity 25.3%; Pred. No. 0.046;  
Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;

QY 10 LQYAEIPKSEIPK-----VIEKAYIPVETLKEIPFGINITYTLKFLPKDII DL 61  
DB 130 MTYVFKESGFPKNRVIGSGVLDTR----FRTFVAEE-----LNIS-----VKDVTGF 175  
QY 62 VKGIASDLIEIGTSYTHAI--LPILPSRVEAQVQDRVEKEELFEVSPKG--FWLPE 117  
DB 176 VLGGHGDWVPLVNRYSVAGIPLKELIPKRLDAIVERTRKGGGEIVNLGNSAYYAPA 235  
QY 118 LAYDPIIPALIKDN-----GYEYLF 137  
DB 236 ASLVEMVEALIKQRRILPALVLEGEYEGEIV 269

## RESULT 12

US-09-540-236-3128  
Sequence 3128, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA.  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 3128  
LENGTH: 653  
TYPE: PR  
ORGANISM: M. catarrhalis  
US-09-540-236-3128

Query Match 5.2%; Score 97; DB 4; Length 653;  
Best Local Similarity 21.9%; Pred. No. 0.37;  
Matches 72; Conservative 43; Mismatches 128; Indels 86; Gaps 13;

QY 22 KYIEKAYIPVETLKEIPFGINITYTLKFLPKDII DL 68  
DB 95 KALIEPRYRGVITD--REGAPLAANAPLYTFDFPAVAYEYRLDNEIKTKSETAKOKA 152  
QY 69 -----DLIEIGTSYTHAILPILPSRVEAQVQDR-----REVEKEELFEVSPKG 113  
DB 153 LKKLKMEDVLRL-----AAANYPLEKLEAAVIGIDHTLDTNSQKKEAL-----PKGA 201  
QY 114 WLPELA-YDPIIPALIKDNNGEYELFADGEAMLFSAHLNSAIKPIKLYPHILKAQREKRF 172  
DB 202 SSRRLVVLNRRVSPVAKSVTDLGLFALIGREOYFORYLOA--EPNAQLGVAQSSDETQ 260  
QY 173 RYISYLLGLRELKAIKIVFSGKVTLKAVKDIEAVPVVAANTVMTGIGLPLMPKRV 232  
DB 261 GYIG-----RAGIEAKYERLAGKDK-----VOILKGTGQPIQELIQI 300  
QY 223 ASWIEDKONILLGYTDIEFIDYRDYAGYRMSVEGLLEVIDELNSHLCLPSELKHSRELY 292  
DB 301 EPLIEGENIRLILDSRLQYVLYKE-----LEQVRLQ-----SARSSSGMVD 343  
QY 293 LRT-----SSMAPDKSLRIRMBEGNAR 315  
DB 344 VKTGEVLTAMGSPSPFNNSNLSERDGNAR 372

## RESULT 13

US-09-134-001C-3678  
Sequence 3678, Application US/09134001C  
Patent No. 6580370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964



; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 3678  
 ; LENGTH: 367  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-3678

Query Match 5.0%; Score 94.5; DB 3; Length 367;  
 Best Local Similarity 18.7%; Pred. No. 0.27;  
 Matches 56; Conservative 52; Mismatches 118; Indels 73; Gaps 12;

QY 4 LVFIGNLOYAIPKSEIPKVEKAYIPYETILKEIPFGNITGYTLKPLKIDLVK 63  
 DB 68 LIFKG-----VRIVEDGGSIIIRKLIONNI--NLIALHTNLDVNPCKGNRMIA 114  
 QY 64 GGIASDLIEIGT--SYTHAILPLPLSRVEAQQVORDREVBELFEVSPKGFMLPELAYD 121  
 DB 115 DQIGENISMINTSSYYKQVTRIPKAYIE-----DFKDSLNEL----- 154  
 QY 122 PIPAILKONGEXYLFADGAMLSAHLNSAIKPIKPLYPHLIKAQREKFRYISYLLGL 181  
 DB 155 ---GLAKGNYECFFESG-----KGFPGVGDASPIGLKDS-----IEYV--- 194  
 QY 182 RELKAIKLVF---EKKVTLKAVKDE--AVPY--WVAVNTAVMLGIGRLPLNPKV 232  
 DB 195 ---DEIDLEFMIKONLEITRKAILDHPYETPVPFDIKKKNKESEYLGIGQLNQMT 250  
 QY 233 ASWIEDKNILLYGTDFIGYRDIAGYRMSVEGLLEV-----IDELNSELCLPSELKH 286  
 DB 251 LDFSEVAKKOLNIPSVARYTGHDSPIKKAIVIGSGIGFHYKASQLGADVFTGDIKH 309

## RESULT 14

US-09-710-279-1840  
 ; Sequence 1840; Application US/09710279  
 ; Patent No. 6703492  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMMERLY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: PUS480US  
 ; CURRENT APPLICATION NUMBER: US/09/710,279  
 ; CURRENT FILING DATE: 2000-11-09  
 ; PRIOR APPLICATION NUMBER: 60/164,258  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1840  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 ; OTHER INFORMATION: amino acid sequence  
 US-09-710-279-1840

Query Match 4.9%; Score 92; DB 4; Length 375;  
 Best Local Similarity 19.3%; Pred. No. 0.51;  
 Matches 72; Conservative 55; Mismatches 109; Indels 138; Gaps 16;

QY 59 IDLVKGIASDLIEIGTSYTHAILPLPLSRV---EAQQVORDREVBELFEVSPK--- 111  
 DB 3 IDISG---DFIAFIGTSGSGKTTALMINRMIESTGEITIDGKNIKEINPVELARSI 58  
 QY 112 GFWEPELAYD-----IIPAILK-----DNGEYV----- 136  
 DB 59 GYVIOQIGLMPHMYKENIVVPPKILKRSQEKDEKAKELIRLVLPPEYIDRYPSSELG 118  
 QY 137 ---FADGAMLSAHLNSAIKPI--KPLYPHLIK---AQREKFRYISYLLGL 181  
 DB 119 GQQRIGVVRALAAEQDIIILMDEPFGALDPIITROTLDLVKQLQQQLKTIPIVTH--- 174

QY 182 RELKAIKLT-----VFEGKVTLKAVKD 203  
 DB 175 -DMDEAIKLDKICIMTNGQVITYDTPDNILRSFANDVRFIQGNRIQRPNIRYKD 233  
 QY 204 IEAVPVVAVTAVMLGIGRLPLNPKVASWIBDKNILLGYTDIEFIGRDIAGYRMS 263  
 DB 234 AMIRPVYVHDSRLNDAVN---IMREKV-----DTIPVGNDEHLGLDIEDINEG 283  
 QY 264 VEGLEVIDELNSEL---LPSELKHSGRELYRTSSWAP-----DKSL----- 304  
 DB 284 LRHHKELIDTQORDIYRVRIKSDSVRTIKGNVRNVPVDSNDKTLGLVTRANLV 343  
 QY 305 ---RIVRE--DEGN 313  
 DB 344 IYDSDIWGELESNG 357

## RESULT 15

US-09-248-796A-15394  
 ; Sequence 15394; Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kelch Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC;  
 ; FILE REFERENCE: 107136.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; CURRENT FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; PRIOR FILING DATE: 1998-08-13  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 15394  
 ; LENGTH: 845  
 ; TYPE: PRT  
 ; ORGANISM: Candida albicans  
 US-09-248-796A-15394

Query Match 4.9%; Score 92; DB 4; Length 845;  
 Best Local Similarity 20.2%; Pred. No. 1.9;  
 Matches 78; Conservative 64; Mismatches 117; Indels 128; Gaps 21;

QY 38 EIDP-FGLNITGYTLKFLPDDIIDLKVGIASDLIEITL-----GTSYTHAI 82  
 DB 202 DEMPHIGVDING-----KRIMPAKGSALDQLLESIDLPEGWTLGLDQWTGS----- 249  
 QY 83 LPLPLSRVEAQQVORDREVBELFE--VSPKGFMLPELAYDPIIPAILKONGEXYLFADGE 141  
 DB 250 ---LKLTDDELELRKIQOQENTDENINP-----YELIDWPTKD----- 286  
 QY 142 AMLPSAHLNSAIKPIKPLYPHLIKAQREKFRYISYLLGRELKAIKLVPEGVK----- 196  
 DB 287 -----EIMP-----VNAVPEPKRRFPSPGHEAKRVKVIKAIAREGIIIPNK 329  
 QY 197 ---TKAVKDIKAVPVM---VAVNTAVM--LGIGRLP-----LNNPKKVA 233  
 DB 330 VKQQLTSEEBEDQNFQMDQDEIRISDHIMWLRAPKLPPTNEESVNPPEEYLLTBEKS 389  
 QY 234 SWIE---DKDNILLYGTDFIEFIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKISGR 289  
 DB 390 KWLQESPIDERNFL---POKYNSLRQVPGQDSVRRF---ERSIDLVLAPVRHN-- 440  
 QY 290 ELVYRTSSWAPD-----KSLR-----IWRDEGNARLNMVLSYNGGELAFILAENG-- 335  
 DB 441 KLNIDPSLIPDLPSPDLLRPPFRCTIYEGHTG--KIRTISIDPOG--LMLATGSDDG 496  
 QY 336 -ARGMEPLPERRLDAPRAIYNDMGEN 361  
 DB 497 SVRIWEILTGROYVYKQILINKEINNEED 523

Search completed: January 27, 2005, 06:00:12  
Job time : 28 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 28, 2005, 04:41:08 ; Search time 573 Seconds

(without alignments)  
3650.093 Million cell updates/sec

Title: US-09-886-400A-4

Perfect score: 1879  
Sequence: 1 LRALVFHGNQYVAIPKSEI.....RLDAPRAIYNDKMGNGEP 364

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1p  
-Q=/cgn2\_1/USPTO.spool.p/US09886400/runat\_27012005\_055724\_17979/app.query.fasta\_1.519  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62  
-TRANS=human40.csl -LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09886400@cgn\_1\_1\_723@runat\_27012005\_055724\_17979  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	1095	15 US-10-112-231A-3	Sequence 3, Appl1
2	1870	99.5	1095	9 US-09-886-400-3	Sequence 3, Appl1
3	1870	99.5	1095	13 US-10-112-357-3	Sequence 3, Appl1
4	1870	99.5	1095	13 US-10-114-403-3	Sequence 3, Appl1
5	1870	99.5	1095	13 US-10-116-606-3	Sequence 3, Appl1
6	1870	99.5	1095	13 US-10-112-331-3	Sequence 3, Appl1
7	1870	99.5	1095	13 US-10-112-377-3	Sequence 3, Appl1
8	1870	99.5	1095	13 US-10-116-581-3	Sequence 3, Appl1
9	1870	99.5	1095	13 US-10-112-442-3	Sequence 3, Appl1
10	1870	99.5	1095	13 US-10-112-418-3	Sequence 3, Appl1
11	1870	99.5	1095	13 US-10-114-083-3	Sequence 3, Appl1
12	111.5	5.9	2670	16 US-10-282-122A-17097	Sequence 17097, A
13	109	5.8	939	15 US-10-369-493-46924	Sequence 46924, A
14	108	5.7	2455	16 US-10-425-114-29659	Sequence 29659, A
15	108	5.7	2586	16 US-10-424-599-32675	Sequence 32675, A
16	106	5.6	1770	18 US-10-425-115-106717	Sequence 106717, A
17	105	5.6	822	9 US-09-974-300-56	Sequence 56, Appl1
18	103	5.5	3222	9 US-09-815-242-8263	Sequence 8263, Ap
19	103	5.5	3441	9 US-09-815-242-4413	Sequence 4413, Ap
20	103	5.5	3453	16 US-10-282-122A-8207	Sequence 8207, Ap
21	103	5.5	5030	8 US-08-781-986A-324	Sequence 324, App
22	103	5.5	5030	16 US-10-329-624-324	Sequence 324, App
23	102.5	5.5	1434	15 US-10-369-493-28193	Sequence 28193, A
24	102.5	5.5	1434	15 US-10-369-493-30951	Sequence 30951, A
25	102.5	5.5	3036	16 US-10-398-221-2003	Sequence 2003, Ap
26	101	5.4	1326	17 US-10-437-963-20338	Sequence 20338, A
27	101	5.4	4755	16 US-10-398-221-3741	Sequence 3741, Ap
28	100.5	5.3	945	15 US-10-369-493-41082	Sequence 41082, A
29	100	5.3	1345	16 US-10-425-114-32190	Sequence 32190, A
30	100	5.3	1441	18 US-10-425-115-136289	Sequence 136289, A
31	99.5	5.3	994	16 US-10-425-114-5764	Sequence 5764, Ap
32	99.5	5.3	994	16 US-10-425-114-5766	Sequence 5766, Ap
33	99	5.2	536165	10 US-09-939-964-1	Sequence 1, Appl1
34	98.5	5.2	1800	15 US-10-369-493-45612	Sequence 45612, A
35	98.5	5.2	1800	18 US-10-793-639-209	Sequence 209, App
36	98.5	5.2	2442	16 US-10-282-122A-11024	Sequence 11024, A
37	98.5	5.2	3082	13 US-10-113-852A-1	Sequence 1, Appl1
38	98	5.2	1533	15 US-10-369-493-23743	Sequence 23743, A
39	97.5	5.2	1342	18 US-10-425-115-51870	Sequence 51870, A
40	97.5	5.2	1590	15 US-10-369-493-33955	Sequence 33955, A
41	97	5.2	1955	16 US-10-282-122A-26871	Sequence 26871, A
42	97	5.2	2925	15 US-10-094-749-581	Sequence 581, App
43	97	5.2	100848	16 US-10-672-787-39	Sequence 39, Appl1
44	96.5	5.1	2498	16 US-10-425-114-20468	Sequence 20468, A
45	96.5	5.1	2914	18 US-10-739-930-1249	Sequence 1249, Ap

#### ALIGNMENTS

RESULT 1  
US-10-112-231A-3  
; Sequence 3, Application US/10112231A  
; Publication No. US20030184276A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR  
; FILE REFERENCE: 09010-004005  
; CURRENT APPLICATION NUMBER: US/10/112,231A  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: US 09/407,806  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 06/613,220  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1095  
; TYPE: DNA

```

; ORGANISM: Thermococcus alcaliphilus
US-10-112-231A-3

Alignment Scores:
Pred. No.: 2,276-209 Length: 1095
Score: 1879.00 Matches: 364
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-886-400A-4 (1-364) x US-10-112-231A-3 (1-1095)

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db 1 TTGAGAGGCGCTCGCTTTCACGGCAACCTCCAGTATGCCAATATCCAAAGAGCGAATC 60

QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db 61 CCAAGGCTCATAGAGGACATACATCCAGTATCCAGACACTGATTAAGAAATTT 120

QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db 121 CTTTGGGCTCAACATACGAGGCTATACCTTAAAGTTCCTCCGAGATATTATTAGAC 180

QY 61 LeuValLysGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db 181 CTGTTAAAGGGGCAATCGGAGATGACTGATAGATATCCGAAGACACTACGCGAC 240

QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
Db 241 GCATATCTCCCTCCCTCCGCTTAGCAGATGAGACCAAGTTCCAGAGATAGAGAA 300

QY 101 ValLysGluGluLeuPheGluValIleSerProLysGlyPheThrLeuProGluLeuAlaTyr 120
Db 301 GTTAAAGAAAGCTCTTCGAGTTTCTCCAAAGGAAATTCGCTGCCAGGCTCGCTTAT 360

QY 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
Db 361 GACCCGATATCTCTGCTGATGAGAGCAACGGTTATGATATCTATTCGCGCAGCGG 420

QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160
Db 421 GAGCGATGTCTTCTTCAGCTCATCTCACTCGCGCATTAAGCCAACTTAACCGCTCTAT 480

QY 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
Db 481 CCACACCTTATTAAGGCCCAAGGAAAGCGCTTATGATCATCATCTATCTCTTGGAT 540

QY 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
Db 541 CTCAGGAGGCTTAGGAAGCGATTAAGCTCGTTTGAAGGTAAAGGTAAACGCTTAAGCA 600

QY 201 ValLysAspIleGluAlaValProValITProValAlaValAsnThrAlaValMetLeuGly 220
Db 601 GTCAAAAGACATCAAAACCGTACCCGTTGGGTGGCCGTAAACACGCTGTATATGCTCGCG 660

QY 221 IleGlyArgLeuProLeuMetAsnProLysValAlaSerThrIleGluLysAsp 240
Db 661 ATCGGAAGCTTCTCTTATGATCTTAAGAAAGTCCAGACTGATAGAGCAAGCAAGGAC 720

QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
Db 721 AACATCTCTTATACCGGACCGCATATAGATTCAATGGCTATAGGAGCACTTCAGAGCTAC 780

QY 261 ArgMetSerValGluLysLeuGluValIleAspGluLeuAsnSerGluLeuLysLeu 280
Db 781 AGAATAGAGTTGAGGATTAATTAGAGTTATAGACAGACTCACTCGGACCTGCTT 840

QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerThrAlaPro 300
Db 841 CCTCAGAGCTGAGACACAGTGAAGGAGCTCTACTTACGAGCTTCGAGTTGGGACCA 900

```

```

QY 301 AspLysSerLeuArgIleThrPargIuAspGluIleAsnAlaArgLeuAsnMetLeuSer 320
Db 901 GATTAAGGCTTGAGGATATGAGAGGAGCAAGGAAAGCAAGCTTATATATGCTGTC 960

QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgIleTyrGlu 340
Db 961 TACAAATATGAGGGGCCAACTCCCTTTTATGCCAGAACAGCGATGCAAGGGGATGGAG 1020

QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspThrArgGlyGlu 360
Db 1021 CCCCTCCCTGAGAGGAGGCTGATGCTCCGGGCGATATATACGATTGAGGGGTGA 1080

QY 361 AsnGlyGluPro 364
Db 1081 AATGGGAACT 1092

RESULT 2
US-09-886-400-3
; Sequence 3, Application US/09886400
; Patent No. US20020045226A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1092)
US-09-886-400-3

Alignment Scores:
Pred. No.: 2,566-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 9 Gaps: 0

US-09-886-400A-4 (1-364) x US-09-886-400-3 (1-1095)

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db 1 TTGAGAGGCGCTCGCTTTCACGGCAACCTCCAGTATGCCAATATCCAAAGAGCGAATC 60

QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db 61 CCAAGGCTCATAGAGGACATACATCCAGTATCCAGACACTGATTAAGAAATTT 120

QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db 121 CTTTGGGCTCAACATACGAGGCTATACCTTAAAGTTCCTCCGAGATATTATTAGAC 180

QY 61 LeuValLysGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db 181 CTGTTAAAGGGGCAATCGGAGATGACTGATAGATATCCGAAGACACTACGCGAC 240

QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100

```

Db 241 GCAATACTCCCTCCCTCCCTAGAGATAGAGCAAGATTGAGAGATAGGAA 300  
Qy 101 VallysglulgluPhegluValSerProlysglyPheTrpLeuProgluValaTy 120  
Db 301 GTTAAGGAGAGCTCTTGAGCTTCTCCAAAGGATTCGCTGCCAGACTCCCTAT 360  
Qy 121 AspProileileProalaileuValaspaanglyTyrgluTyLeuPhealaaspGly 140  
Db 361 GACCGATATATCCCTGCACTACTGAAAGCAAGGATATGATATCTATTCGCGAGCGG 420  
Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlailelyProilelyProleuTy 160  
Db 421 GAGGCGAGCTTTTCTGAGCTCATCTCAACTCGGAGATTAAGCCGATTAACCGCTAT 480  
Qy 161 ProHisLeuilelyValaGlnaArgGluValArgPheArgTyHiserTyLeuLeuGly 180  
Db 481 CCACACCTTATTAAGGCCCAAGGAAAGCGCTTAGGTATCATCATCTCTCTGCT 540  
Qy 181 LeuArgGluLeuArglyValailelyValaValPhegluGlyValValThreuleuVala 200  
Db 541 CTCAGGAGCTTGAAGAGCCATTAAGCTCGTTTGAAGGTAAAGGTAAAGGCA 600  
Qy 201 VallysApillegluValaValProValTrpValaValaAsnThrAlaValMetLeuGly 220  
Db 601 GTCAAGACATCGAAGCCGTACCGCTTGGGTGGCCGTGAACAGCGCTGATATGCTCGC 660  
Qy 221 IleGlyArgLeuProleuMetAsnProlyValaValaserTrpIleGluAspLyasp 240  
Db 661 ATCGAAGGCTTCTCTTATATATCTTAAGAAAGTGGAGCTGATAGAGACAAAGAC 720  
Qy 241 AsnIleleuLeuTyrglyTyThrAspIleGluPheileGlyTyArgAspIlealaGlyTy 260  
Db 721 AACATCTTCTTAACGGCAAGATATAGATTCATTCGCTATAGGACATTCAGAGCTAC 780  
Qy 261 ArgMetSerValGluGlyLeuLeuGluValileAspGluLeuAsnSerGluLeuCyseu 280  
Db 781 AGAATGAGGTGTGAGGAGATTTATAGAGTTATAGCGAGCTCACTCGGACTGTGCTT 840  
Qy 281 ProSerGluLeuValHisSerGlyArgGluLeuTyLeuArgThrSerSerTrpAlaPro 300  
Db 841 CCTCAGAGCTGAAGACAGTGAAGGAGGAGCTCTTACGAGACTTCGAGTGGGCAACA 900  
Qy 301 AspySerLeuValgileTrpArgGluAspGluValaValaValaValaValaVala 320  
Db 901 GATAGAGCTTGAAGATATGAGAGAGCAAGAGGAAAGCAAGCTTAATATGCTGTCC 960  
Qy 321 TyraMetArgGlyGluLeuAlaPheleuAlaGluAsnSerAspAlaArgGlyTrpGlu 340  
Db 961 TACAATATGAGGCGCAACTCGCTTTTACCGCAAGAACAGCATGCAAGGAGATGGAG 1020  
Qy 341 ProleuProgluArgArgLeuAspAlaPheArgAlaileTyraAspTrpArgGlyGlu 360  
Db 1021 CCCCTCCCTGAGAGAGGCTGATGCTTCCTCGGCGATATATTAAGATGAGGGGTAA 1080  
Qy 361 AsnGlyGluPro 364  
Db 1081 AATGGGGAACCT 1092

## RESULT 3

US-10-112-357-3  
; Sequence 3, Application US/10112357  
; Publication No. US20020115099A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: Murphy, Dennis  
; APPLICANT: Ried, John  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH  
; FILE REFERENCE: DIVER1120-4  
; CURRENT APPLICATION NUMBER: US/10/112.357  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 09/886,400  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 09/619,032

; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/407,806  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 08/613,220  
; PRIOR FILING DATE: 1996-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: Thermococcus alcaliphilus  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1092)  
US-10-112-357-3  
Alignment Scores:  
Pred. No.: 2,56e-208 Length: 1095  
Score: 1870.00 Matches: 362  
Percent Similarity: 99.73% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.52% Indels: 0  
Gaps: 0  
US-09-886-400a-4 (1-364) x US-10-112-357-3 (1-1095)  
Qy 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyraAlaGluileProlySerGluile 20  
Db 1 TTGAGAGCGCTCGCTTTCACGGCACTCCAGATATCCGAATCCCAAGCGAAATC 60  
Qy 21 ProlyValileGlyValaTyTrileProValileGluTrpLeuilelyGluGluile 40  
Db 61 CCAAGGATGATGAGAGGAGCATATACATCCAGTCAGACATGATTAAGAGAAATTT 120  
Qy 41 ProbheGlyLeuAsnileThnGlyTyThrLeuysPheLeuProlyAspIleileasp 60  
Db 121 CTTTGGGCTCAACATACACGGGCTATACCTTAAGTTCTCCCGGAAGATATATAC 180  
Qy 61 LeuValysgllyileleAsnSerAspLeuileGluileileGlyThrSerTyThrHis 80  
Db 181 CTGTTAAAGGGGATCGGAGATGATCTGATAGATTAATCGGAAGAGCTACACGAC 240  
Qy 81 AlaileleuProleuLeuProleuSerArgValGluAlaGluValaGlnaArgAspGlu 100  
Db 241 GCAATACTCCCTCCCTCCCTGCTTGAAGAGTAAACAAAGTTCAAGAGATAGGAA 300  
Qy 101 VallysglulgluPhegluValSerProlysglyPheTrpLeuProgluValaTy 120  
Db 301 GTTAAGGAGAGCTCTTGAGCTTCTCCAAAGGATTCGCTGCCAGACTCCCTAT 360  
Qy 121 AspProileileProalaileuValaspaanglyTyrgluTyLeuPhealaaspGly 140  
Db 361 GACCGATATATCCCTGCACTACTGAAAGCAAGGATATGATATCTATTCGCGAGCGG 420  
Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlailelyProilelyProleuTy 160  
Db 421 GAGGCGAGCTTTTCTGAGCTCATCTCAACTCGGAGATTAAGCCGATTAACCGCTAT 480  
Qy 161 ProHisLeuilelyValaGlnaArgGluValArgPheArgTyHiserTyLeuLeuGly 180  
Db 481 CCACACCTTATTAAGGCCCAAGGAAAGCGCTTAGGTATCATCATCTCTCTGCT 540  
Qy 181 LeuArgGluLeuArglyValailelyValaValPhegluGlyValValThreuleuVala 200  
Db 541 CTCAGGAGCTTGAAGAGCCATTAAGCTCGTTTGAAGGTAAAGGTAAAGGCA 600  
Qy 201 VallysApillegluValaValProValTrpValaValaAsnThrAlaValMetLeuGly 220  
Db 601 GTCAAGACATCGAAGCCGTACCGCTTGGGTGGCCGTGAACAGCGCTGATATGCTCGC 660  
Qy 221 IleGlyArgLeuProleuMetAsnProlyValaValaserTrpIleGluAspLyasp 240  
Db 661 ATCGAAGGCTTCTCTTATATATCTTAAGAAAGTGGAGCTGATAGAGACAAAGAC 720

QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260  
 DB 721 AACATCTCTTATACGACGACGATATAGATTTCATGGCTATAGGACATTGCAGGCTAC 780  
 QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280  
 DB 781 AGAATGAGTGTGAGGCGATTATTAGAGGTATTAGCGAGCTCAACTCGGAACTGTGCTT 840  
 QY 281 ProSerGluLeuLeuHisSerGlyArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300  
 DB 841 CCTCAGAGCTGAGACACAGTGAAGGAGCTTACCTTACCGACTTCGAGTGGGCACCA 900  
 QY 301 AspLysSerLeuArgIleTyrArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320  
 DB 901 GATAGAGCTTGAAGATATGAGAGAGACGAAAGGAAACGCAACATTAAATATCTGTCC 960  
 QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340  
 DB 961 TACAAATATGAGGGCGCAACTGCCCTTTTACCGAAGACGATGCAAGGCGATGGAG 1020  
 QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360  
 DB 1021 CCCCCTCCCTGAGAGAGGCTGATGCTTCGCGGCGATATATACGATTGGAGGGGTGAA 1080  
 QY 361 AsnGlyGluPro 364  
 DB 1081 AATGGGGAACCT 1092

## RESULT 4

US-10-114-403-3  
 ; Sequence 3, Application US/10114403  
 ; Publication No. US20020115100A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/114,403  
 ; PRIOR FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: 09/886,400  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR APPLICATION NUMBER: 08/613,220  
 ; PRIOR FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1095  
 ; TYPE: DNA  
 ; ORGANISM: Thermococcus alcaliphilus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1092)  
 ; US-10-114-403-3

## Alignment Scores:

Pred. No.: 2,566-208 Length: 1095  
 Score: 1870.00 Matches: 362  
 Percent Similarity: 99.73% Conservative: 1  
 Best Local Similarity: 99.45% Mismatches: 1  
 Query Match: 99.52% Indels: 0  
 DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-114-403-3 (1-1095)

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1081 AATGGGGAACCT 1092

DB 1 TTGAGACCGCTCGCTTTTACGCGCAACCTCCAGTATGCCGAATCCCAAGACGCAATC 60  
 QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluIle 40  
 DB 61 CCAAGGCTCATAGAGAGGATATCATCCAGTATGAGACACTGATTAAAGAAATTT 120  
 QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60  
 DB 121 CTTTTGGGCTCAACATAACGGGCTATTACCTTAAAGTTCTCCCGAAGATATTATAGAC 180  
 QY 61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80  
 DB 181 CTGCTTAAAGGGGCGATCCGAGTGCCTGATAGATATATGGAACGAGCTTACACGAC 240  
 QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluValGlnArgAspArgGlu 100  
 DB 241 GCATATCTCCCTCTCCCGCTTAGCAGATAGAAAGCAAGATTGAGAGATAGGGA 300  
 QY 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTyrLeuProGluLeuAlaTyr 120  
 DB 301 GTTAAAGAAAGCTCTTCAGAGCTTCTCCAAAGGGAATTCGCTGCGCAGAGCTCGCTAT 360  
 QY 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140  
 DB 361 GACCCGATATATCCCTGCCATCTGAAAGCAACGTTATGATATCTATTCGCCGACGG 420  
 QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160  
 DB 421 GAGGCGATGCTTTTCTCAGCTCATCTCAACTCGGGGATAAACCAATTAAACGCTCAT 480  
 QY 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180  
 DB 481 CCACACTTATTAAGGCCCAAGAGGAAAGCCCTTTAGTATCATCAGCTATCTCTTGCT 540  
 QY 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200  
 DB 541 CTCAGGAGCTTAGAAGGCCATTAAGCTCGTTTGAAGGTAAAGCTTAAGGCA 600  
 QY 201 ValLysAspIleGluAlaValProValIleValAlaValAsnThrAlaValMetLeuGly 220  
 DB 601 GTCAAGAATCAGGAAGCCGTACCCGTTGGGGCGCGGAACAACGCTGTAATGCTCGGC 660  
 QY 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspLysAsp 240  
 DB 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGAGCTGATAGAGGACAAAGAC 720  
 QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260  
 DB 721 AACATCTCTTATACGACCGGATATAGATTTCATGGCTATAGGACATTGCAGGCTAC 780  
 QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280  
 DB 781 AGAATGAGTGTGAGGAGTATTATAGAGGTATTAGCGAGCTCAACTCGAAGCTGCGCTT 840  
 QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300  
 DB 841 CCTCAGAGCTGAGACACAGTGAAGGAGCTTACCTTACCGACTTCGAGTGGGCACCA 900  
 QY 301 AspLysSerLeuArgIleTyrArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320  
 DB 901 GATAGAGCTTGAAGATATGAGAGAGACGAAAGGAAACGCAACATTAAATATCTGTCC 960  
 QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340  
 DB 961 TACAAATATGAGGGCGCAACTGCCCTTTTACCGAAGACGATGCAAGGCGATGGAG 1020  
 QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360  
 DB 1021 CCCCCTCCCTGAGAGAGGCTGATGCTTCGCGGCGATATATACGATTGGAGGGGTGAA 1080  
 QY 361 AsnGlyGluPro 364  
 DB 1081 AATGGGGAACCT 1092



LOCATION: (1) ... (1092)  
US-10-112-331-3

Alignment Scores:  
Pred. No.: 2,566-208 Length: 1095  
Score: 1870.00 Matches: 362  
Percent Similarity: 99.73% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.52% Indels: 0  
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-112-331-3 (1-1095)

```

QY      1 LeuAArgAlaLeuValPheHisGlyAenLeuGlnTyrAlaGluIleProLysSerGluIle 20
DB      1 TTGAAGAGCGCTCGCTTTTCACGGCAACCTCCAGATAGCCGAATCCCAAGAGCGCAATC 60

QY      21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
DB      61 CCAAGGCTCATAGAGAGGCAATACATCCAGTCATCGAGACATGATTTAAAGAAATTT 120

QY      41 ProPheGlyLeuAenIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
DB      121 CTTTTGGGCTCAACATTAACGGGCTATACCTTAAAGTTCCCTCCGAAGATATTATAGAC 180

QY      61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
DB      181 CTGGTTAAAGGGGCGATCGAGAGTACCTGATAGATTAATCGAAGCGACTACAGCAC 240

QY      81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
DB      241 GCATATACCTCCCTCCCTCCCGCTTAGAGAGTAGAACAATGTCAGAGATAGAGGA 300

QY      101 ValLysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTyr 120
DB      301 GTTAAAGAGAGGCTCTTCGAGCTTTCTCCAAAGGAGATTCTGGCTGCCAGAGCTCGCTAT 360

QY      121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrTrpLeuPheAlaAspGly 140
DB      361 GAACCCGATTAATCCCTCGCATACGAAAGACAACGGTATAGATTCATTCGCGCAGAG 420

QY      141 GluAlaMetLeuPheSerAlaHisLeuAenSerAlaIleLysProIleLysProLeuTyr 160
DB      421 GAGGCGATGCTTTTCTCAGCTCATCTCACTCGGCGAATTAAGCAATTAACGCTTAT 480

QY      161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
DB      481 CCACACCTTAAAGGCCCAAGGGAAGACGCTTTAGTACATCAAGCTATCTCTTGCT 540

QY      181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
DB      541 CTGAGGAGCTTGAAGAGCGAATTAACCTGTTTTTAAGAGTAAAGGTAAGCTTAAGGCA 600

QY      201 ValLysAspIleGluAlaValProValTrpValAlaValAsnThrAlaValMetLeuGly 220
DB      601 GTCAAAACATTCGAAGCCGTAACCGCTTGGGTGGCCGCTGAACACGCTGTAATGCTCGC 660

QY      221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTrpIleGluLysPlyAsp 240
DB      661 ATCGGAAGGCTTCTCTTATGAATCTTAAGAAAGTGGCAGCTGGAATGAGGACAGAC 720

QY      241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
DB      721 AACATTTCTTAACGGCACCGAATAGAGTCAATTTGCTTAAGGAGCATTTGCAAGGCTAC 780

QY      261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAenSerGluLeuCysLeu 280
DB      781 AGATAGAGTGTGAGGATTAATTAAGGTTATTAACAAGCTCAACTCGGAACGTGCTT 840

QY      281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300
DB      841 CCTCAAGAGCTGAAGCACAGTGAAGGAGACTTACTTACGACCTTGAAGTTGGGACCA 900

```

```

QY      301 AspLysSerLeuArgIleTrpArgGluAspGluGlyAsnAlaArgLeuAenMetLeuSer 320
DB      901 GATTAAGCTTAGAGATATAGAGAGACGAAGGAAAGCAAGACTTATATATGCTGTCC 960

QY      321 TyrAenMetArgGlyGluLeuAlaPheLeuAlaGluAenSerAspAlaArgGlyTrpGlu 340
DB      961 TCAATATATAGAGGCGGAACTCCCTTTTACCGGAGAACGAGATGCCAAGGGGATGGAG 1020

QY      341 ProLeuProGluLysArgLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
DB      1021 CCCCTCCCTGAAGAGAGGCTGAGTGCCTTCGAGGATATTAACGATTCGAGGGGTGA 1080

QY      361 AsnGlyGluPro 364
DB      1081 AATGGGAACT 1092

```

# RESULT 7

```

US-10-112-377-3
; Sequence 3, Application US/10112377
; Publication No. US20020120108A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,377
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alkaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1092)
; US-10-112-377-3

```

Alignment Scores:  
Pred. No.: 2,566-208 Length: 1095  
Score: 1870.00 Matches: 362  
Percent Similarity: 99.73% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.52% Indels: 0  
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-112-377-3 (1-1095)

```

QY      1 LeuAArgAlaLeuValPheHisGlyAenLeuGlnTyrAlaGluIleProLysSerGluIle 20
DB      1 TTGAAGAGCGCTCGCTTTTCACGGCAACCTCCAGATAGCCGAATCCCAAGAGCGCAATC 60

QY      21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
DB      61 CCAAGGCTCATAGAGAGGCAATACATCCAGTCATCGAGACATGATTTAAAGAAATTT 120

QY      41 ProPheGlyLeuAenIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
DB      121 CTTTTGGGCTCAACATTAACGGGCTATACCTTAAAGTTCCCTCCGAAGATATTATAGAC 180

QY      61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
DB      181 CTGGTTAAAGGGGCGATCGAGAGTACCTGATAGATTAATCGAAGCGACTACAGCAC 240

```



Qy 81 AlaileuProleuLeuProleuSerArgValGluValGlnValGlnArgAspArgGlu 100  
Db 241 GCAATACCTCCCTCTCTCCGCTTAGCAGAGTGAAGACCAAGTTCAAGAGATGGGAA 300  
Qy 101 VallysglulLeuPheGluValSerProlysglyPheTrpLeuProGluLeuValTyr 120  
Db 301 GTTAAGGAAGAGCTCTTCGAGCTTCTCCAAAGGAGATTCTGGCTCCAGAGCTCCCTAT 360  
Qy 121 AspProileileProAlaileuLeuAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140  
Db 361 GACCGAATATCCCTGCAATCTGAAGACCAAGGTTATGATCTATTCGCGCAAGG 420  
Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaileuPheProileuPheProleuTyr 160  
Db 421 GAGGCGATGCTTTTCTCAGCTCATCTCACTCGGAGATTAACCAATTAAACGCTCTAT 480  
Qy 161 ProHileuileuValaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180  
Db 481 CCACACCTTATTAAGGCCCAAGGAAAGCCCTTAGGTACATCAGCTATCTCTTGGT 540  
Qy 181 LeuArgGluLeuValaGlyValaileuValaileuValaPheGluGlyLysValaThrLeuVala 200  
Db 541 CTCAGGAGCTTATGAGGAAGCGATTAAGCTCGTTTGAAGGTAAAGTAAACGCTAAAGCA 600  
Qy 201 VallysAspIleGluAlaValaProValaTrpValaValaAsnThrAlaValaMetLeuGly 220  
Db 601 GTCAAGACATCGAAGCCGTAACCTTTGGGTGGCGGTGAACAGCGCTGTAATGCTTCGAC 660  
Qy 221 IleGlyArgLeuProleuMetAsnProlysglyValaIleSerTrpIleGluAspLysAsp 240  
Db 661 ATCGGAAGCTTCTCTTATGATCTTAAGAAAGTGGGAGCTGATGAAGACAAAGAC 720  
Qy 241 AsnIleleuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260  
Db 721 AACATCTCTTATACGGCAGATTAAGATTGCTATGAGGACATTCAGGCTAC 780  
Qy 261 ArgMetSerValaGluGlyLeuLeuGluValaIleAspGluLeuAsnSerGluLeuCysLeu 280  
Db 781 AGAATGATGTTGAGGAGTTTATTAAGGTTTATAGCGAGCTCACTCGAAGCTGGCTT 840  
Qy 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300  
Db 841 CCTCAGAGCTGAGACCAAGTGAAGGAGCTTACTTACGACCTTCAGATTGGGACCA 900  
Qy 301 AspLysSerLeuArgIleTyrArgGluAspGlyLysAsnAlaArgLeuAsnMetLeuSer 320  
Db 901 GATAAGAGCTTGAAGATTAAGAGAGGACGAGGAGGAAAGCAAGCTTAATATGCTGTCC 960  
Qy 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTrpGlu 340  
Db 961 TACAATATGAGGAGGAGGAACTGCTTTTACCGGAAACAGGATGCAAGGAGGAGAG 1020  
Qy 341 ProleuProGluArgArgLeuAspAlaPheArgAlaileuTyrAsnAspTrpArgGlyGlu 360  
Db 1021 CCCCTCCCTGAGAGGAGCTGATCCCTTCGGGCGGATATTAAGATTGAGGGGTAA 1080  
Qy 361 AsnGlyGluPro 364  
Db 1081 AATGGGAACT 1092

## RESULT 8

US-10-116-581-3  
; Sequence 3, Application US/10116581  
; Publication No. US20020137116A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: Murphy, Dennis  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
; FILE REFERENCE: DIVER1120-4  
; CURRENT APPLICATION NUMBER: US/10/116,581  
; CURRENT FILING DATE: 2001-04-03

; PRIOR APPLICATION NUMBER: 09/886,400  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 09/619,032  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/407,806  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 08/613,220  
; PRIOR FILING DATE: 1996-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: Thermococcus alcaliphilus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1) ... (1092)  
US-10-116-581-3

## Alignment Scores:

Pred. No.: 2,566-208 Length: 1095  
Score: 1870.00 Matches: 362  
Percent Similarity: 99.73% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.52% Indels: 0  
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-116-581-3 (1-1095)

Qy 1 LeuArgAlaLeuValaPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluLe 20  
Db 1 TTGAAGAGCTTCGCTTTCACGAGCACTCCAGTATCCGAAATCCCAAGAGGAAATC 60  
Qy 21 ProLysValaIleGluValaTyrIleProValaIleGluThrLeuIleLysGluGluLe 40  
Db 61 CCAAGGCTATAGAGGAAGCATATCATCCAGTCAAGACATGATTAAGAAAGAAATT 120  
Qy 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleLeuAsp 60  
Db 121 CCTTTGGGCTCAACATTAACGGGCTATACCTTAAGTTCTCCGAAAGGATTAATATGAC 180  
Qy 61 LeuValaLysGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80  
Db 181 CTCGTTAAAGGGGAGATCGAGATGACTGATTAAGATTAATCGAGATCAAGCAC 240  
Qy 81 AlaileuProleuLeuProleuSerArgValaGluValaGlnValaGlnArgAspArgGlu 100  
Db 241 GCAATACCTCCCTCTCTCCGCTTAGCAGAGTGAAGACCAAGTTCAAGAGATGGGAA 300  
Qy 101 VallysglulLeuPheGluValaSerProlysglyPheTrpLeuProGluLeuValTyr 120  
Db 301 GTTAAGGAAGAGCTTCTTCAGCTTCTCCAAAGGAGATTCTGGCTCCAGAGCTCCCTAT 360  
Qy 121 AspProileileProAlaileuLeuAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140  
Db 361 GACCGAATATCCCTGCAATCTGAAGACCAAGGTTATGATCTATTCGCGCAAGG 420  
Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaileuPheProileuPheProleuTyr 160  
Db 421 GAGGCGATGCTTTTCTCAGCTCATCTCACTCGGAGATTAACCAATTAAACGCTCTAT 480  
Qy 161 ProHileuileuValaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180  
Db 481 CCACACCTTATTAAGGCCCAAGGAAAGCCCTTAGGTACATCAGCTATCTCTTGGT 540  
Qy 181 LeuArgGluLeuValaGlyValaileuValaileuValaPheGluGlyLysValaThrLeuVala 200  
Db 541 CTCAGGAGCTTATGAGGAAGCGATTAAGCTCGTTTGAAGGTAAAGTAAACGCTAAAGCA 600  
Qy 201 VallysAspIleGluAlaValaProValaTrpValaValaAsnThrAlaValaMetLeuGly 220  
Db 601 GTCAAGACATCGAAGCCGTAACCTTTGGGTGGCGGTGAACAGCGCTGTAATGCTTCGAC 660

QY 221 ILeGlyArGLeuProLeuMeLeAsnProLySylsValAlaSerTriPleGluAspLyAsp 240  
Db 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGCACTGATAGAGCAAGGAC 720  
QY 241 AsnIleLeuLeuTyrgLyThrAspIleGluPheIleGlyTyraAspIleAlaGlyTyr 260  
Db 721 AACATTCTTCATACCGGACCGCATATAGATTCAATTGGCTATAGGACATTCGAGGCTAC 780  
QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280  
Db 781 AGAATGAGTGTGGAGGATTTATAGAGTTATAGACAGCTCACTGGAATGTGTCTT 840  
QY 281 ProSerGluLeuTyrsSerGlyArGlyLeuTyrgLyLeuArGlyThrSerSerTriPalaPro 300  
Db 841 CCTCAGAGCTGAAGCAAGTGAAGGAGCTCACTTACGGACTTGAGTTGGGACCA 900  
QY 301 AspLySerSerLeuArGlyIleTPArGlyuAspGluGlyAsnAlaArGLeuAsnMetLeuSer 320  
Db 901 GATTAAGAGCTTGAAGATATGAGAGAGAGCAAGGAGCAAGCACTTAATATGTCTTCC 960  
QY 321 TyrAsnMetArGlyGlyLeuLeuAlaPheLeuAlaGluAsnSerAspAlaArGlyTTPGlu 340  
Db 961 TACAAATGAGGGGCGAACTCGCCCTTTAGCGAGAACAGCGATGCAAGGGATGGAG 1020  
QY 341 ProLeuProGluArGlyLeuAspAlaPheArGlyAlaIleTyraAsnAspTriPArGlyGlu 360  
Db 1021 CCCCTCCCTGAGAGAGAGCTGATGCTTCCGGGCGATATATAAGATTGGAGGGGTGA 1080  
QY 361 AsnGlyGluPro 364  
Db 1081 AATGGGAACTT 1092  
RESULT 9  
US-10-112-442-3  
; Sequence 3, Application US/10112442  
; Publication No. US20020150997A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: Murphy, Dennis  
; APPLICANT: Ried, John  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH  
; FILE REFERENCE: DIVER1120-4  
; CURRENT APPLICATION NUMBER: US/10/112,442  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 09/886,400  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 09/619,032  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/407,806  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 08/613,220  
; PRIOR FILING DATE: 1996-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: Thermococcus alcaliphilus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1092)  
US-10-112-442-3  
Alignment Scores:  
Pred. No.: 2,566-208 Length: 1095  
Score: 1870.00 Matches: 362  
Percent Similarity: 99.73% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.52% Indels: 0  
DB: 13 Gaps: 0  
US-09-886-400A-4 (1-364) x US-10-112-442-3 (1-1095)

QY 1 LeuArGAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLySerGluIle 20  
Db 1 TTGAGAGCGCTCGCTCTTTACCGGCAACTCCAGTATGCCGAATCCCAAGAGCGAAATC 60  
QY 21 ProLyValAlleGluTyrsAlaTyrIleProValIleGluThrLeuIleGlyGluIle 40  
Db 61 CCAAGGTCAATAGAGAGCAATACATCCAGTCATGAGACACTGATTAAAGAAATTT 120  
QY 41 ProPheGlyLeuAsnIleThrGlyTyrgLyThrLeuTyrgLyPheLeuProLyAspIleLeaP 60  
Db 121 CCTTTGGGCTCAACATTAACGGGCTATACCTTAAGTTCTCCCGGAAGATATTATAGAC 180  
QY 61 LeuValIleGlyGlyIleAlaSerAspLeuIleGlyIleGlyTyrgLyThrHis 80  
Db 181 CTGTTAAAGGGGGCAATCCGAGTACCTGATAGATATATCGAAACGAGCTACAGGCAC 240  
QY 81 AlaIleLeuProLeuLeuProLeuSerArGlyAlaGluAlaGluValGlnArGAspArGly 100  
Db 241 GCAATACTCCCTCTCTCCGCTTACAGATAGAAAGCAAGTTCAAGAGATAGAGAA 300  
QY 101 ValLyGlyGlyLeuPheGlyValSerProLyGlyPheTyrgLyProGluLeuAlaTyr 120  
Db 301 GTTAAGGAAGAGCTTTCGAGCTTCTCCAAAGGAGTTCTGGCTGCCAGACTCGCTAT 360  
QY 121 AppProIleIleProAlaIleLeuTyrsAspAsnGlyTyrGlyTyrLeuPheAlaAspGly 140  
Db 361 GACCCGATATATCCCTGCGCATCTGAAAGCAACGGTTAGATATCTATTCGCCAGCGGG 420  
QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleTyrgLyProIleTyrgLyProLeuTyrg 160  
Db 421 GAGGCGATGCTTTTCTGAGCTCATCTCAACTGGGGAATAAACCAATTAAACGCTCAT 480  
QY 161 ProHisLeuIleTyrgLyAlaGlnArGlyTyrgLyPheArGlyTyrIleSerTyrgLeuGly 180  
Db 481 CCACACTTATATAAGCCCAAGGAGAAAGCCCTTTAGGTATCATCATCTCTTGCT 540  
QY 181 LeuArGlyLeuArGlyValIleTyrgLyValIlePheGlyTyrgLyValThrLeuTyrsAla 200  
Db 541 CTCAGGAGCTTATAGAAAGCGCATTAAGCTCGTTTGAAGGATTAACGCTTAAGGGA 600  
QY 201 ValLyAspIleGlyAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220  
Db 601 GTCAAAAGCATGAGAGCGGTACCGCTTGGGGCGCGTGAACAGCGCTGATATGCTCGGC 660  
QY 221 ILeGlyArGLeuProLeuMeLeAsnProLySylsValAlaSerTriPleGluAspLyAsp 240  
Db 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGCACTGATAGAGCAAGGAC 720  
QY 241 AsnIleLeuLeuTyrgLyThrAspIleGluPheIleGlyTyraAspIleAlaGlyTyr 260  
Db 721 AACATTCTTCATACCGGACCGCATATAGATTCAATTGGCTATAGGACATTCGAGGCTAC 780  
QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280  
Db 781 AGAATGAGTGTGGAGATTAATAGAGTTATAGAGTTCAGAGCTCAACTCGGAATGTGTCTT 840  
QY 281 ProSerGluLeuTyrsSerGlyArGlyLeuTyrgLyLeuArGlyThrSerSerTriPalaPro 300  
Db 841 CCTCAGAGCTGAAGCAAGTGAAGGAGCTCACTTACGGACTTGAGTTGGGACCA 900  
QY 301 AspLySerSerLeuArGlyIleTPArGlyuAspGluGlyAsnAlaArGLeuAsnMetLeuSer 320  
Db 901 GATTAAGAGCTTGAAGATATGAGAGAGAGCAAGGAGCAAGCACTTAATATGTCTTCC 960  
QY 321 TyrAsnMetArGlyGlyLeuLeuAlaPheLeuAlaGluAsnSerAspAlaArGlyTTPGlu 340  
Db 961 TACAAATGAGGGGCGAACTCGCCCTTTAGCGAGAACAGCGATGCAAGGGATGGAG 1020  
QY 341 ProLeuProGluArGlyLeuAspAlaPheArGlyAlaIleTyraAsnAspTriPArGlyGlu 360  
Db 1021 CCCCTCCCTGAGAGAGAGCTGATGCTTCCGGGCGATATATAAGATTGGAGGGGTGA 1080  
QY 361 AsnGlyGluPro 364

Db 1081 AATGGGAACCT 1092

## RESULT 10

US-10-112-418-3

```

Sequence 3, Application US/10112418
Publication No. US20020155486A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
APPLICANT: Ried, John
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/112,418
CURRENT FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: 09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1095
TYPE: DNA
ORGANISM: Thermococcus alkaliphilus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1092)
US-10-112-418-3

```

## Alignment Scores:

```

Pred. No.: 2,566-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

```

US-09-886-400a-4 (1-364) x US-10-112-418-3 (1-1095)

```

Qy 1 LeuAArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db 1 TTGAGACGCGCTCGCTTTCACGCGCAACCTCCAGTATGCCGAATCCCAAGAGCGAATC 60
Qy 21 ProLysValIleGluValAlaTyrIleProValIleGluThrLeuIleLysGluIle 40
Db 61 CCAAGGTCATAGAGAGGCAATACCTCCATCCAGACACTGATTAAGAAAGAAATT 120
Qy 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db 121 CTTTTGGGCTCAACATAGCGGCTATACCTTAAGTTCCCTCCGGAAGATATTATGAC 180
Qy 61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db 181 CTCCTTAAGGGGCGCATCGCAGTGACCTGATAGATTAATCGGAACGAGCTACACGAC 240
Qy 81 AlaIleLeuProLeuLeuProLeuSerArgValGluIleValGlnValGlnAspArgGlu 100
Db 241 GCATTAATCTCCCTCCCTCCGCTTAGCGAGATAGAGCAAAAGTTCCAGAGATAGGAAA 300
Qy 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTyrPleuProGluLeuAlaTyr 120
Db 301 GTTAAGAGAGGCTCTTGAGCTTTCTCCAAAGGATTTCTGGCTGCCAGAGCTGCTAT 360
Qy 121 AspProIleIlePheProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
Db 361 GACCCGATTAATCCCTCGCATCTAGAGAGCAACGTTATGATTTATTCGCCAGCGGG 420

```

```

Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160
Db 421 GAGCGATGCTTTTCTAGCTCATCTCACTCGCGGAGTAAGCCATTAAACCGCTCAT 480
Qy 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLysGly 180
Db 481 CCACACCTTAAGAGCCCAAGGAAAGCGCTTAGTGTACATCAGCTATCTCCTGGT 540
Qy 181 LeuArgGluLeuAspGlyValAlaIleLysLeuValPheGluGlyLysValThrLeuValAla 200
Db 541 CTCAGGAGCTTAGAGGCGATTAAGCTCTCTTTTAAGAGTAAGTAACCTTAAGGCA 600
Qy 201 ValLysAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
Db 601 GTCAAGAGCTTCAGAGCCGTAACCTGTTGGGTGGCCGTGAACAGCGCTGTAATCTCGGC 660
Qy 221 IleGlyValLeuProLeuMetAsnProLysValAlaSerTyrIleGluAspLysAsp 240
Db 661 ATCGAAGGCTTCTCTTATGATATCTTAAGAAAGTGCGAGCTGATAGAGCAAGGAC 720
Qy 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
Db 721 AACATTTCTTATACGGCACCGATATAGATTCAATTGGCTATAGGACATTGCAAGCTAC 780
Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
Db 781 AGAATGATGTTGGAGATTTTAAAGCTTTAACAAGCTCACTCGAATCTGCTT 840
Qy 281 ProSerGluLeuLysHisSerGlyValArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300
Db 841 CCTCAGAGCTGAGACAGTGAAGGAGAGCTTACTTACGACTTGAGTGGAGCACCA 900
Qy 301 AspLysSerLeuAlaGlyIleTyrArgGluAspGluLysAsnAlaArgLeuAsnMetLeuSer 320
Db 901 GATTAAGAGCTTAGAGATATGAGAGAGAGACGAAAGGAACTTAATATCTGTGCC 960
Qy 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
Db 961 TACATATTGAGGCGCAACTCGCTTTTACCCAGAAACCGATGCAAGGAGATGGAG 1020
Qy 341 ProLeuProGluArgTyrLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360
Db 1021 CCCCTCCCTGAGAGAGGCTGAGATGCTTCCGGGCGATATATTAACGATTGGAGGGGTGAA 1080
Qy 361 AsnGlyGluPro 364
Db 1081 AATGGGAACCT 1092

```

## RESULT 11

US-10-114-083-3

```

Sequence 3, Application US/10114083
Publication No. US20020160464A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
APPLICANT: Ried, John
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/114,083
CURRENT FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1095

```

```

; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1092)
US-10-114-083-3

```

```

Alignment Scores:
Pred. No.: 2,56e-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

```

```

US-09-886-400a-4 (1-364) x US-10-114-083-3 (1-1095)

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProlysSerGluIle 20
DB 1 TTGAGAGCGCTGCTCTTTTCACGGCAACCTCCAGTATGCCAAATCCCAAGAGCGCAATC 60

QY 21 ProlyValIleGluIysAlaTyrIleProValIleGluThrLeuIleGluIle 40
DB 61 CCAAGGTCATAGAGGAGCATACATCCAGTCATCGACACTGATTAAAGAAAT 120

QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuIlePheLeuProlysAspIleIleAsp 60
DB 121 CTTTGGGCTCAACATTAAGGCTATACCTTAAAGTCTCCCGAAGGATATTATAGAC 180

QY 61 LeuValIleGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
DB 181 CTGTTAAAGGGGCGATCGGAGTACCTGATAGAGTATCCGAAGAGACTACACGCAC 240

QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluValGluArgAspArgGlu 100
DB 241 GCATACTCCCGCTCCCGCTTACGAGTAGAGCAAGATTCCAGAGATAGGAA 300

QY 101 ValIleGluIleLeuPheGluValSerProlyGlyPheThrLeuProGluLeuAlaTyr 120
DB 301 GTTAAAGAGAGCTCTTGGAGCTTCTCCAAAGGATTTGGCTGCCAGAGCTGCTAT 360

QY 121 AspProIleIleProAlaIleLeuIleAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
DB 361 GACCCGATATCCCTGCGATACGAAAGACACGTTATGATATCTATTCGCGAGCGG 420

QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleIleProIleIleProLeuTyr 160
DB 421 GAGGCGATGCTTTCTCAGCTCATCTCAACTCGCGCATTAAGCCAAATTAACCGCTAT 480

QY 161 ProHisLeuIleIleValIleGluArgGluIlePheArgTyrIleSerTyrLeuLeuGly 180
DB 481 CCACACCTTAAAGGCCCAAGGAAAGGCGTTTGGTACATCACTATCTCTTGT 540

QY 181 LeuArgIleLeuArgIleValIleLeuValPheGluIleGlyValThrLeuIleVala 200
DB 541 CTCGAGGAGCTTAGAGAGCGCATTAACCTCTTTTGAAGTAAAGTAAAGCTAAAGCA 600

QY 201 ValIleAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
DB 601 GTCAAAACATCGAAGCGCTACCGCTTGGGTGCGCTGAAACAGCGCTGTAATCTCGGC 660

QY 221 IleGlyArgLeuProLeuMetAspProlyIleValIleAspTyrIleGluAspAsp 240
DB 661 ATCGAAGGCTTCTCTTATATATCTTAAGAAATGGGAGCTGATAGAGAAAGAGAC 720

QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
DB 721 AACATTCTTCTATACGGACCGATATAGATTCTTGGCTATAGAGCATTCGAGGCTAC 780

QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
DB 781 AGAATGAGGTGTGAGGAGTATATAGAGGTATAGACGAGCTCACTCGGAATCTGCTT 840

```

```

QY 281 ProSerGluLeuIleSerGlyArgGluLeuTyrIleArgThrSerSerTrpAlaPro 300
DB 841 CCTTAGAGCTGACACAGTGAAGGAGCTCTTACTTACGAGCTTCAGATTGGGCACCA 900

QY 301 AspIleSerLeuArgIleTrpArgIleAspGluIleAsnAlaArgLeuAsnMetLeuSer 320
DB 901 GATTAAGAGCTTGAAGATATGAGAGAGACGAAAGGAAACGCAAGACTTAATATGCTGTCC 960

QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
DB 961 TGCATATAGAGGGGCGAACTGCTCTTTAGCCGAACAGCGATGCAGAGGAGTGGAG 1020

QY 341 ProLeuProGluIleArgGlyLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
DB 1021 CCCCTCCCTGAGAGGAGGCTGATGCTTCCGGCGATATATACGATTGAGGGGTGA 1080

QY 361 AsnGlyGluPro 364
DB 1081 AATGGGGAACCT 1092

RESULT 12
US-10-282-122A-17097
; Sequence 17097, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELIIRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17097
; LENGTH: 2670
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-282-122A-17097

Alignment Scores:
Pred. No.: 0.0446 Length: 2670
Score: 111.50 Matches: 72

```



```

US-10-425-114-29659
; Sequence 29659, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Yindong
; APPLICANT: Zhou, Yinhua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29659
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GWFLMINSOY109F02_FLI
US-10-425-114-29659

Alignment Scores:
Pred. No.: 0.101 Length: 2455
Score: 108.00 Matches: 85
Percent Similarity: 36.73% Conservative: 59
Best Local Similarity: 21.68% Mismatches: 106
Query Match: 5.75% Indels: 142
DB: 16 Gaps: 21

US-09-886-400a-4 (1-364) x US-10-425-114-29659 (1-2455)

QY 6 PheHisGlyAsnLeu---GlnTyrAlaGluIleProLysSer-----Glu 19
DB 1415 TTTCTGGGATCAGTCTGGAAAGCTTGGTGGTTCATCATCACTCGGGGCTGAGTGGGAG 1474
QY 20 IleProLysValIleGluLysAlaTyrIleProVal----- 31
DB 1475 GTCCCTCCCAATGGTGCACCCATTTTCTCCATATTATTGGTGTCCCGCAGGTATT 1534
QY 32 -----IleGluThrLeu 35
DB 1535 TCTACCTGCTCCTTCATGTGACGAGCATTCACCAATCCCAACATCATATTGAAACATG 1594
QY 36 IleLysGluGluIleProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuPro 55
DB 1595 -----CCTTTCCCTCAGGCTCTTATTATGACCAATCCTCATG----- 1636
QY 56 LysAspIleIleAspLeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGly 75
DB 1637 -----GTAACTGTGG-----GATCAGTACAACTCTCGGT 1669
QY 76 ThisSerTyrThrHisAlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluVal 95
DB 1670 ACTTCATGAGCACTTCTCT-----CCTTTTTCCT-----GATCAGTGTGTC 1711
QY 96 GlnArgAspArgGluValLysGluGluLeuPheGluValSerProLysGlyPheTyrPheLeu 115
DB 1712 AGAATGCTCTTGGCAACATCTCAGCAATTCACCAATTCACCA----- 1756
QY 116 ProGluLeuAlaLysPheProIle-----IleProAlaIle-----LeuLysAspAsnGly 132
DB 1757 -----TTAATGTGATTCATATGTTTCATGTCCAGTGTGATGACGCTCTTCAGGT 1810
QY 133 TyrGluTyrLeuPheAlaAspGlyGluAlaMetLeuPheSerAlaHisLeuAsnSerAla 152
DB 1811 CAGGGCTATCTTGTAGGTGTGGCCCTGCAATGTCACCTTAC----- 1852
QY 153 IleLysProIleLysProLeuTyrProHisLeuIleLysAlaGlnArgGluLysArgPhe 172
DB 1853 -----ATTCACCATGATGATCAACACCTTGTGAGGACATGATTCTGTGAATGTGAT 1903

```

```

QY 173 ArgTyrIleSerTyrLeuLeuGlyLeuArgGluLeuArgLysAlaIleLysLeuValPhe 192
DB 1904 GCTGTGTGTG-----AAAGGTGCAAGAGAGACTCTAAGATTGCTCAT 1945
QY 193 GluGly-----LysValThrLeuLysAlaValLysAsp 203
DB 1946 AGTGGTTCAAGCCAGGAGTAAACCAAGATGATGAGGACACTTGCCTGCATTT----- 1999
QY 204 IleGluAlaValProValIleTyrValAlaValAsnThrAlaValMetLeuGlyIleArg 223
DB 1999 ----- 1999
QY 224 LeuProLeuMetAsnProLysLysValAlaIleSerTyrIleGluAspLysAspAsnIleLeu 243
DB 2000 -----TTAAGTAAATCAAC-----GAAATCAAAACATATTCCT 2035
QY 244 LeuTyrGlyTyrAspIleGluPheIleGlyTyrArgAspIle----- 257
DB 2036 GTGGCTGGAAGTGTGTGCTTTCACACTGACCAAGATATTATCTATTGCAACAGC 2095
QY 258 -----AlaGlyTyrArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsn 275
DB 2096 ATGCTGCTATGGAGATTAGATGATCGTCTGGGGGTATCCAGGTACAGGTGAGCTAT 2155
QY 276 SerGluLeuGlyLeuProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThr 295
DB 2156 TCGGACTATATGT-----GAAACTATGGAATTTGGACACAGCAAGAAAT 2200
QY 296 SerSerTyrAlaProAspLysSerLeuArgIleTyrArgGluAspGluGlyLysAlaArg 315
DB 2201 TCCAAC-----GATTCAGGTGTGTGCTTTTGGATGATGAGGAGGATCTTCT 2251
QY 316 Leu-----AsnMetLeuSerTyrAsnMetArgGly 325
DB 2252 TTGAGTTTGAAGAGAGTGTAGTATGATGAATGCAATGTCTTTTCTTTTATTATGTA 2311
QY 326 GluLeuAlaPheLeuAlaGluAsnSerAspAlaArg 337
DB 2312 GATCTCTTTTGTATTACCA---AATGTATGATATGCCG 2344

RESULT 15
US-10-424-599-32675
; Sequence 32675, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 28564
; SEQ ID NO 32675
; LENGTH: 2586
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12950C.1
US-10-424-599-32675

Alignment Scores:
Pred. No.: 0.109 Length: 2586
Score: 108.00 Matches: 85
Percent Similarity: 36.73% Conservative: 59
Best Local Similarity: 21.68% Mismatches: 106
Query Match: 5.75% Indels: 142
DB: 16 Gaps: 21

US-09-886-400a-4 (1-364) x US-10-424-599-32675 (1-2586)

```



***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 28, 2005, 02:20:58 ; Search time 109 Seconds

(without alignments)  
2373.644 Million cell updates/sec

Title: US-09-886-400A-4

Perfect score: 1879  
Sequence: 1 LRALVPHGNLQYAEIPKSEI.....RLDAFPAIYNDMGNGECP 364

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n\_model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_epool\_p/US09886400/runat\_27012005\_055722\_17947/app\_query.fasta\_1.519  
-DB=Issued\_Patents\_NA -QEMT=fastap -SUFFIX=rnl -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09886400@cgn2\_1\_105@runat\_27012005\_055722\_17947 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	1095	4	US-10-166-606-3
2	1879	100.0	1095	4	US-10-112-231A-3
3	1589	84.6	1041	2	US-08-613-220B-3
4	146.5	7.8	1664976	4	US-08-916-421B-1
5	146.5	7.8	1664976	4	US-09-692-570-1
6	140.5	7.5	3139	1	US-07-894-212A-1
7	105	5.6	684	4	US-09-107-532A-1900
8	104	5.5	4403765	3	US-09-103-840A-2
9	104	5.5	4411529	3	US-09-103-840A-2
10	103	5.5	5030	4	US-08-956-171E-324
11	103	5.5	5030	4	US-08-956-171E-324
12	102.5	5.5	1912	1	US-08-270-013B-1

13	102.5	5.5	1912	1	US-08-838-418-1	Sequence 1, Appli
14	101.5	5.4	1083	4	US-09-710-279-2149	Sequence 2149, Ap
15	101.5	5.4	1131	4	US-09-710-279-2105	Sequence 2105, Ap
16	101.5	5.4	3014	4	US-09-710-279-4330	Sequence 4330, Ap
17	101.5	5.4	3041	4	US-09-710-279-4336	Sequence 4368, Ap
18	99	5.2	536165	4	US-09-214-808-1	Sequence 1, Appli
19	98.5	5.2	1800	4	US-09-614-221A-209	Sequence 209, App
20	97	5.2	1662	4	US-09-540-236-1208	Sequence 1208, App
21	97	5.2	100848	4	US-09-596-002-199	Sequence 39, Appli
22	94.5	5.0	1104	3	US-09-134-001C-841	Sequence 841, App
23	94.5	5.0	3183	4	US-09-221-017B-939	Sequence 939, App
24	93.5	5.0	100990	4	US-09-409-800B-2	Sequence 2, Appli
25	92	4.9	1128	4	US-09-710-279-1839	Sequence 1839, Ap
26	92	4.9	1664976	4	US-09-252-991A-15134	Sequence 15134, A
27	92	4.9	2226	4	US-09-252-991A-14877	Sequence 14877, A
28	92	4.9	2268	4	US-09-252-991A-14887	Sequence 14887, A
29	92	4.9	2538	4	US-09-248-796A-1291	Sequence 1291, Ap
30	92	4.9	4086	4	US-09-710-279-1765	Sequence 3765, Ap
31	92	4.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
32	92	4.9	1664976	4	US-09-692-570-1	Sequence 1, Appli
33	91.5	4.9	3030	4	US-09-693-146-3	Sequence 3, Appli
34	91.5	4.9	3113	4	US-09-894-998A-52	Sequence 52, Appli
35	91.5	4.9	3345	4	US-09-894-998A-49	Sequence 49, Appli
36	91.5	4.9	3350	4	US-09-894-998A-48	Sequence 48, Appli
37	91.5	4.9	5421	1	US-08-118-441-28	Sequence 28, Appli
38	91.5	4.9	5421	3	US-08-338-579A-28	Sequence 28, Appli
39	91.5	4.9	5421	5	PCT-US94-09851-28	Sequence 28, Appli
40	91	4.8	3096	4	US-09-543-681A-4073	Sequence 4073, App
41	90.5	4.8	1221	3	US-09-134-001C-733	Sequence 733, App
42	90.5	4.8	2370	4	US-09-248-796A-5191	Sequence 5191, Ap
43	90.5	4.8	3516	4	US-09-248-796A-1940	Sequence 1940, Ap
44	90	4.8	4074	4	US-09-252-991A-4737	Sequence 4737, Ap
45	90	4.8	4092	4	US-09-252-991A-4771	Sequence 4771, Ap

#### ALIGNMENTS

RESULT 1  
US-10-166-606-3  
; Sequence 3, Application US/10166606  
; Patent No. 6644756  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR  
; FILE REFERENCE: 09010-004005  
; CURRENT APPLICATION NUMBER: US/10/166, 606  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: US 09/407, 806  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 08/613, 220  
; PRIOR FILING DATE: 1996-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: Thermococcus alcaliphilus  
US-10-166-606-3

#### Alignment Scores:

Pred. No.: 1,166-230 Length: 1095  
Score: 1879.00 Matches: 364  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-886-400A-4 (1-364) x US-10-166-606-3 (1-1095)

0y 1 LeuArgLauValPheHisGlyAsnLeuGlnTyrLacGluIleProLysSerGluIle 20  
|||||

Db 1 TTGAGAGCGCTGCTTTTCAGCGCAACCTCCAGTATGCGAAATCCCAAGAGCGAAATC 60  
 Qy 21 ProlyValIleGluValAlaTyrIleProValIleGluThrLeuIleValGluGluIle 40  
 Db 61 CCANAGTCAATAGAGAGCATACATCCAGTCAATCGAGACACGATTTAAAGAAATTT 120  
 Qy 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuValPheLeuProLysAspIleGluAsp 60  
 Db 121 CCTTTGGGCTCAACATAGAGGGCATACCTTAAAGTTCCCTCCGAAAGATTTATAGAC 180  
 Qy 61 LeuValIysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80  
 Db 181 CTCCTTAAAGGGGCATCGCAGTGCATGATAGATTAATCGGAACGAGCTACGCCAC 240  
 Qy 81 AlaIleLeuProLeuLeuProLeuSerArgValGluIleGluValGluIleAlaArgAspArgGlu 100  
 Db 241 GCAATATCTCCCTCTCTGCGCTTACGAGTATGACACACAAAGTTGAGAGATAGAGAA 300  
 Qy 101 ValIysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTyr 120  
 Db 301 GTTAAAGAGAGCTCTTCGAGGTTTCTCCAAAGGGAATTCGGCTGCCAGAGCTCGCTAT 360  
 Qy 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140  
 Db 361 GACCCGATATCTCTGCGCATACGTAAGACAAACGGTATGATATCTATTCGCCGACGG 420  
 Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160  
 Db 421 GAGCGATGCTTTTCTGAGCTCATCTCACTCGCGATTAAGCCAAATTAACGCTCTAT 480  
 Qy 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180  
 Db 481 CCACACCTTAAAGGCCCAAGGAAAGCGCTTATGATACATCAGCTATCTCTTGCT 540  
 Qy 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200  
 Db 541 CTCAGGAGCTTGAAGAGCGATTAAGCTGTTTTGAAAGTAAAGTAAAGCTAAAGCGCA 600  
 Qy 201 ValIysAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220  
 Db 601 GTCAAGACATCGAAGCGCATACCGTTGGGTGGCGGTGAACCGGCTGTAAGGCTGGC 660  
 Qy 221 IleGlyArgLeuProLeuMetAsnProLysValAlaSerTrpIleGluAspLysAsp 240  
 Db 661 ATGGAGAGCTTCTCTTATGAATCCTAAGAAAGTGGCGAGTGAATAGAGCAAGAC 720  
 Qy 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260  
 Db 721 AACATTTCTTATACGACCGCATATAGAGTTCAATGGCTATAGGACATTTGACAGCTAC 780  
 Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCySerLeu 280  
 Db 781 AGATGAGGTGAGGGAATTATAGAGTTATAGAGCTCACTCGAATCTGCTCTT 840  
 Qy 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300  
 Db 841 CCTTCAGAGCTGAGACACGTGAGAGGAGCTCTACTTACGAGCTCGAGTTGGCACCA 900  
 Qy 301 AspLysSerLeuArgIleTyrArgGluAspGluGluValAsnAlaArgLeuAsnMetLeuSer 320  
 Db 901 GATTAAGCTTGAGGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Qy 321 TyrAsnMetArgGlyGlyLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340  
 Db 961 TACAAATATAGAGGCGCACTCGCTTTTACCGGACACACGAGATCCAAAGGAGATGGAG 1020  
 Qy 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360  
 Db 1021 CCCCTCCCTGAGAGGAGGCTGATGCTTCCGGGCGATATATACATTTGAGAGGAGTGA 1080  
 Qy 361 AsnGlyGluPro 364  
 Db 1081 AATGGGAACT 1092

RESULT 2  
 US-10-112-231A-3  
 ; Sequence 3, Application US/10112231A  
 ; Patent No. 6744246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Reid, John  
 ; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR  
 ; FILE OF INVENTION: MAKING AND USING THEM (Amended)  
 ; FILE REFERENCE: 09010-004005  
 ; CURRENT APPLICATION NUMBER: US/10/112,231A  
 ; PRIORITY FILING DATE: 2002-03-29  
 ; PRIORITY FILING DATE: US 09/407,806  
 ; PRIORITY FILING DATE: 1999-09-28  
 ; PRIORITY FILING DATE: 1996-03-08  
 ; PRIORITY FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1095  
 ; TYPE: DNA  
 ; ORGANISM: Thermococcus alkaliphilus  
 US-10-112-231A-3

## Alignment Scores:

Pred. No.: 1,166-230 Length: 1095  
 Score: 1879.00 Matches: 364  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-886-400A-4 (1-364) x US-10-112-231A-3 (1-1095)

Qy 1 LeuArgAlaLeuValPheHisLeuGluGlyTyrAlaGluIleProLysSerGluIle 20  
 Db 1 TTGAGAGCGCTGCTTTTCAGCGCAACCTCCAGTATGCGAAATCCCAAGAGCGAAATC 60  
 Qy 21 ProlyValIleGluValAlaTyrIleProValIleGluThrLeuIleValGluGluIle 40  
 Db 61 CCANAGTCAATAGAGAGCATACATCCAGTCAATCGAGACACGATTTAAAGAAATTT 120  
 Qy 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuValPheLeuProLysAspIleGluAsp 60  
 Db 121 CCTTTGGGCTCAACATAGAGGGCATACCTTAAAGTTCCCTCCGAAAGATTTATAGAC 180  
 Qy 61 LeuValIysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80  
 Db 181 CTCCTTAAAGGGGCATCGCAGTGCATGATAGATTAATCGGAACGAGCTACGCCAC 240  
 Qy 81 AlaIleLeuProLeuLeuProLeuSerArgValGluIleGluValGluIleAlaArgAspArgGlu 100  
 Db 241 GCAATATCTCCCTCTCTGCGCTTACGAGTATGACACACAAAGTTGAGAGATAGAGAA 300  
 Qy 101 ValIysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTyr 120  
 Db 301 GTTAAAGAGAGCTCTTCGAGGTTTCTCCAAAGGGAATTCGGCTGCCAGAGCTCGCTAT 360  
 Qy 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140  
 Db 361 GACCCGATATCTCTGCGCATACGTAAGACAAACGGTATGATATCTATTCGCCGACGG 420  
 Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160  
 Db 421 GAGCGATGCTTTTCTGAGCTCATCTCACTCGCGATTAAGCCAAATTAACGCTCTAT 480  
 Qy 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180  
 Db 481 CCACACCTTAAAGGCCCAAGGAAAGCGCTTATGATACATCAGCTATCTCTTGCT 540  
 Qy 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200

```

Db      541 CTCAGGAGCTTAGAGAGCGATTAAGCTGTTTGAAGGTAAAGTAACGCTAAAGCA 600
Qy      201 Vallysaapilleglualavalprovalttrpvalalavalantthralavalmetleuily 220
Db      601 GTCAAGACATCGAAGCCGATACCCCTTTGGGGGCGGAAACAGCGCTGTAATGCTCGGC 660
Qy      221 lleglyargleuproleumetasnprolylsvalalasetripilleglunaplyasap 240
Db      661 ATCGGAAGCTTCCTCTTATGATCTTAAGAAAGTGGGAGCTGATAGAGACAAAGAC 720
Qy      241 AsnilleuleuleuYrGlyThraspilleglubheilleglyYrargaspilleaglYtyr 260
Db      721 AACATTTCTTATACGGCACCGCATATAGATTGCTATAGGACATTCAGGACTAC 780
Qy      261 ArgmetserValgluglYleuleugluvalilleaspgluauasensergluleucyaleu 280
Db      781 AGAATGAGGTGTGAGGAGTATATAGAGGTATATAGCGAGCTCAATCGGAACGTGCGCTT 840
Qy      281 ProsergluleuYrHieSerglyYarggluleuYrleuAqThrSerserTripalapro 300
Db      841 CCTCAGAGCTGAAGACAGTGAAGGAGGCTCTTACGACCTTCAGATTGGGACACA 900
Qy      301 AsplyserleuargiletrparggluapgluglYasnalAargleuasnmetleuser 320
Db      901 GATTAAGAGCTTGAGGATATGAGAGAGGACGAAAGGAGACGCAAGCTTATATGCTGTCC 960
Qy      321 TyrasmelarglygluleuAlaphelenuagluuanserserapalAarglyYtrpilu 340
Db      961 TACAATATGAGGGGCGAACTGCTTTTACCCGAAACACGCAATGCAAGGGGATGGGAG 1020
Qy      341 ProleuprogluAargargleuapalapheargalaileryrasnasptriparglyglu 360
Db      1021 CCCCTCCCTGAGAGAGGCTGATGCCCTCCGGGCGATATATACGATTGAGGGGTGA 1080
Qy      361 AsnlygluPro 364
Db      1081 AATGGGGAACT 1092

```

## RESULT 3

```

US-08-613-220B-3
; Sequence 3, Application US/08613220B
; Patent No. 5958751
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/613,220B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-68-5099
;

```

```

;
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1038
; OTHER INFORMATION:
US-08-613-220B-3

Alignment Scores:
Pred. No.: 1.55e-193 Length: 1041
Score: 1589.00 Matches: 346
Percent Similarity: 95.05% Conservative: 0
Best Local Similarity: 95.05% Mismatches: 0
Query Match: 84.57% Indels: 18
DB: 2 Gaps: 18

US-09-886-400A-4 (1-364) x US-08-613-220B-3 (1-1041)
Qy      1 LeuarglaleuValPheHieglYAsnleuglntYrAlagluileProlySersgluile 20
Db      1 TTGAGAGCGCTCGCTCTTTCACGCGACACTCCAGTATGCCAAATCCCAAGAGCGAA 57
Qy      21 ProlyValillegluYsAlaYrilleProvalillegluThrleuileYsgluile 40
Db      58 CCAAGGTCATAGAGAGGCAATACATCCAGTATCGAGACATGATTAAAGAGAA 114
Qy      41 ProheglYleuAnillethnglyYrThrleuYsPheleupProlySapilleleasp 60
Db      115 CTTTGGGCTCAACATACCGGCTATACCTTAAAGTCTCCCGAAGGATATTATA 171
Qy      61 LeuValysglYglYleAlasersapleuilegluilejleglYThrserYrThris 80
Db      172 CTCGTTAAAGGGGCGATCGGAGTACCTGTATAGATTAATCGAAGAGCTACAG 228
Qy      81 AlalileupProleuProleuSersargValgluAlagluValAlagatgasparglu 100
Db      229 GCATATCTCCCGCTCGCGCGCTTACGAGTGAACACAAAGTTGAGAGATAG 285
Qy      101 VallysleuYrleuPhegluValSerProlysglyPheYrleupProgluYleuAlayr 120
Db      286 GTTAAAGAGGCTCTTCGAGGTTTCCTCAAGGGATTCTGCGTCCAGAGCTGCC 342
Qy      121 AspProillepProAlilleuYsAsnGlyYrGluYrleuPheAlaAspGly 140
Db      343 GACCGATTAATCCCTGCCATATCTGAAGACAAACGGTTATGATTAATTCGCGAC 399
Qy      141 GluAlametleupheseralAlisleuasnserAlilleyspProilleyspProleuYr 160
Db      400 GAGCGATGCTTTTCTAGCTCATCTCACTCGCGATTAAGCAATTAAACCGCTC 456
Qy      161 ProhileuileYsAlagluYrGluYsPheargYrilleSeryrleuYleugly 180
Db      457 CCACACCTTATTAAGGCCCAAGGAAACCGCTTGTGATACATCACTACTCTCT 513
Qy      181 LeuargluYleuYrYsAlaileYsleuValPhegluYleYsValThrleuYsAla 200
Db      514 CTCAGGAGCTTAGAGAGGCGATTAAGCTGTTTGAAGGTAAAGTAAGCTTAAG 570
Qy      201 Vallysaapilleglualavalprovalttrpvalalavalantthralavalmetleugly 220
Db      571 GTCAAGACATCGAAGCCGATACCCGTTGGGTGGCCGTGAACAGGCTGTAATGCTC 627
Qy      221 lleglyargleuproleumetasnprolylsvalalasetripilleglunaplyasap 240
Db      628 ATCGGAAGCTTCCTCTTATGAAATCCTAAGAAAGTGGGAGCTGATAGAGGACAG 684
Qy      241 AsnilleuleuYrGlyThraspilleglubheilleglyYrargaspilleaglYtyr 260

```

Db	665	AACATCTCTTATATAGGCACCGCATTTAAGTTCATTGGCTATAGGCAATTGCAGCC---	741
OY	261	ArgMetSerValGluGluYleuLeuValIleAspGluLeuAsnSerGluLeuGlySerLeu	280
Db	742	AGAAAGATGTTGAGGAGATTATTAGAGTTTATAGACGAGCTCAATCGGAAC TGCC---	798
OY	281	ProSerGluLeuLysHisSerGluYArgGluLeuTyrLeuArgThrSerSerTProAlaPro	300
Db	799	CCCTCAGAGCTGGAACACAGTGGAAGGAGGCTCTCTTACCGAAGCACTTCGAGTTGGCCA---	855
OY	301	AspLysSerLeuArgIleTyrAlaGluLysAspGluLysAlaArgLeuAsnMetLeuSer	320
Db	856	GATAGAGCTTGAGGATATGAGAGAGAGACGAGGAAACGCAAGACTTAATATATGCTG---	912
OY	321	TyrAsnMetArgGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGluTyrProIlu	340
Db	913	TACATATGAGAGGGGGAAGCTCGCTTTTAACTCCGAGAACAGCGATGCAAGGGAGATG---	969
OY	341	ProLeuProGluArgArgLeuAspAlaPheAlaArgAlaIleTyrAsnAspTyrArgGlu	360
Db	970	CCCCCTCCCTGAGAGGAGGCTGAGTGCCTTCGGGGGATATATTAACGATTTGAGGGGGT---	1026
OY	361	AsnGluGluPro	364
Db	1027	AATGGGGAACCT	1038

```

RESULT 4
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PR275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: patentin version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2822)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)

```

OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (98343) ..(98343)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (98343) ..(98343)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (103998) ..(103998)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (148548) ..(148548)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (163385) ..(163385)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (191989) ..(191989)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (191995) ..(191995)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (231580) ..(231580)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (234187) ..(234187)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (234220) ..(234220)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (234814) ..(234814)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (309398) ..(309398)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (309418) ..(309418)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (312837) ..(312837)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (312993) ..(312993)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (319226) ..(319226)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (559167) ..(559167)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (559241) ..(559241)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (600992) ..(600992)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (622708) ..(622708)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (657081) ..(657081)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (657203) ..(657203)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (674435) ..(674435)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (682442) ..(682442)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (713652) ..(713652)
OTHER_INFORMATION: n equals a, t, c, or g	NAME/KEY: misc feature	LOCATION: (713652) ..(713652)

```

NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1115881)..(1115881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1565020)..(1565020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

```

```

Alignment Scores:
Pred. No.: 0.00235 Length: 1664976
Score: 146.50 Matches: 87
Percent Similarity: 40.58% Conservative: 66
Best Local Similarity: 23.08% Mismatches: 129
Query Match: 7.80% Gaps: 95
DB: 4

```

```

US-09-886-400a-4 (1-364) x US-08-916-421B-1 (1-1664976)
QY 14 GluileProlySerGluileProlyValileGluValileProValileGlu 33
DB 1550843 AAATTAATAAGAAAGTTTAAATTAAGTGAATTAATGCTAATTCCTAACAATGAG 1550784

```

```

QY 34 Thrileuileys-----GluGluileProPheGlyLeuAan-----IleThrGly--- 48
DB 1550783 TTGATATTGACGCTTATTGAGTAATATTTAAAGTTATTTATTCATTTACTGGGCT 1550724
QY 49 -----TyrThrleuileysPheleuProlySerPheleuValileGly 64
DB 1550723 TTTGAGAGCAAGCTTTGGAATTT-----AATGATTCAGTTGGATTATTTAAAGAT 1550670
QY 65 GlyIleAlaSerApleuileGluileIleleGlyThrSerTyrThrIleAlaIleleuPro 84
DB 1550669 TTGTAAAACTGGCAATGTGAGTTGATGCTGAAACATATTCATCTCACTCAACAAGT 1550610
QY 85 LeuileuProleu-----SerArgValGluValleGluValleGluValleGluValleGlu 103
DB 1550609 CTATTGTAAGTGAAGATGAGTTATTTAAGATTTGATGATGATGATGATGATGATGATGATGAT 1550550
QY 104 GluileuPheGluValleSerProlyGlyPheTyrleuProGluValleValleValleValle 123
DB 1550549 GAATATTGTTGTTTAAAGCCAGGTTTGAATACTGAGCTTATTAACAACAACA 1550490
QY 124 IleProAlaIleleuValleSerPheGlyTyrGluTyrleuPheAlaSerGly---GluAla 142
DB 1550489 ATTGCAAGATGACAAAGATTTTGGGTTTAAAGCATATTACTGAGGGAATTGAGAAA 1550430
QY 143 MetleuPheSerAlaIleleuAanSerAlaIleleuProIleuProIleuProIleuProIleu 162
DB 1550429 ATCTTA----- 1550424
QY 163 LeuileuValleGluValleGluValleGluValleGluValleGluValleGluValleGlu 179
DB 1550423 -----GGCTGAGGCTCCCAACAGCTCTTATTCATCAACAGATGATGATGATGATGAT 1550373
QY 180 GlyleuValleGluValleGluValleGluValleGluValleGluValleGluValleGlu 199
DB 1550372 CTGTTAAGAAATTTAGTTAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 1550328
QY 200 AlaValleVallePheIleGluAlaVallePro-----ValTyrValleAla 212
DB 1550327 TCAGCAAGATTTGGGACCAATATTCATTAACAGCTGATTAATTCCTATTGTTGTTACT 1550268
QY 213 ValleAanThrAlaValleMetleuGlyIleGlyArgleuProleuMetAanProlyleVal 232
DB 1550267 TCACTCTCGGGAAGTTATTAACATA-----TATATGACATAGACATTT 1550220
QY 233 AlaSer-----TyrIleGluApleuValleAanIleleuLeuTyrGlyThrApleu--- 249
DB 1550219 GGAGAACATCACTGAAAGAA-----ACTGAATATTT 1550187
QY 250 GluPheIleGlyTyrArg-----ApleuIleApleuTyr---ArgMetSerValleGluGly 266
DB 1550186 GAGTTTGTGAGATTTTGGCTTATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1550127
QY 267 LeuileuGluValleAanGluValleAanSerGluValleuProSerGluLeuIleuHis 286
DB 1550126 GTCACTGAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1550094
QY 287 SerGlyArgGluLeuTyrleuArg-----ThrSerTyrPala---ProApleuVal 302
DB 1550093 -----GGAGAGATTTACGTCATGAAATTTGCAATATGATGATGATGATGATGATGAT 1550040
QY 303 SerleuArgIleTyrPalaValleGluValleGluValleGluValleGluValleGluValle 322
DB 1550039 GATGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549992
QY 323 MetArgGlyleuValle-----PheleuAlaGluValleAanSerApleuValleGlyTyr 341
DB 1549991 AAATTAAGATATTTGTAATTTATTAAGAGAACTCAACAATTA----- 1549944
QY 342 LeuProGluValleArgleuAanPalaPheArgAlaIleTyrAanPheTyrArg 358
DB 1549943 -----AAAAAATTAAATTTGATGAAATATATATAGATGATTAAG 1549902

```

RESULT 5

US-09-692-570-1/c  
Sequence 1, Application US/09692570  
Patent No. 6797466  
GENERAL INFORMATION:  
APPLICANT: Bult et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
Patent No. 6797466  
FILE REFERENCE: PB275c1  
CURRENT APPLICATION NUMBER: US/09/692,570  
CURRENT FILING DATE: 2003-01-14  
PRIOR APPLICATION NUMBER: US 60/024,428  
PRIOR FILING DATE: 1996-08-22  
PRIOR APPLICATION NUMBER: US 08/916,421  
PRIOR FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:

```

NAME/KEY: misc feature
LOCATION: (682442) .. (682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (71352) .. (71352)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684) .. (741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455) .. (779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676) .. (779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539) .. (855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619) .. (871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830) .. (1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846) .. (1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119881) .. (119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1130881) .. (1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988) .. (1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224) .. (1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473) .. (1349473)
OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:
Pred. No.: 0.00235 Length: 1664976
Score: 146.50 Matches: 87
Percent Similarity: 40.58% Conservative: 66
Best Local Similarity: 23.08% Mismatches: 129
Query Match: 7.80% Indels: 95
DB: 4 Gaps: 24

US-09-886-400A-4 (1-364) x US-09-692-570-1 (1-1664976)

OY 14 GluilePpolySergluilePpolyValilegnybalyrileProvalileGlu 33
Db 1550843 AAATTAAATAAAGAAGTTTAAATTAAGTGCCTAATAATCTCAATCTCAACAAATGAG 1550784
OY 34 Threuleileys-----GluGluileProheGlyleuan-----llethGly--- 48
Db 1550783 TTGAATATGGAGCTTAATTAATGAATATGATCAATTTAAAGTAAATTAATTAATTCATTAATTAATCTGGAGGTT 1550724

```

49 -----TyrThrLeuValPheLeuProLysAsp---IleIleAspLeuValIleGly 64  
Db 1550723 TTGTGAGGCGACAGCTTGGAATTT-----AATGATTAACGTGGTGAATTTTAAAGAT 1550670  
GlyIleAlaSerAspLeuIleGluIleIleGlyThrSerYThrHisAlaIleLeuPro 84  
Db 1550669 TTGGTAAATACTGGCAATGTTGAGTGTATAGCTGAACATATCATCTCACTCAACAAAGT 1550610  
LeuLeuProLeu---SerArgValGluAlaGlnValGlnArgAspArgGluValIleGlu 103  
Db 1550609 CTATTTGAACCTGAAGATGACTTTATTGAAAGTATTAGATGTCATAAAGAGTTACAAA 1550550  
GluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaIleArgProIle 123  
Db 1550549 GAATATTGTTGTTTAAAGCCAGAGTGTTTAAGAAATCTAGACGTAATATACACACACAGA 1550490  
IleProAlaIleLeuLysAspAsnGlyTYrGluTYrLeuPheAlaAspGly---GluAla 142  
Db 1550489 ATTGCAAAAGATAGCAAAAGATTTTAGGGTTTAAACGATATTACTGAGGGGAATTGAGAA 1550430  
MetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTYrProHis 162  
Db 1550429 ATCTTA----- 1550424  
LeuIleLysAlaGlnArgGluLysArgPheArgTYrIleSer-----TYrLeuLeu 179  
Db 1550423 -----GGCTGAGAGCTCTCCAAACTATCTTTCCAAATCCACAGTGGTATGAAATTT 1550373  
GlyLeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLys 199  
Db 1550372 CTGTTAAGAAATTATAGGTTGAGTGAGTACGATCTGGCTTTGA-----TTT 1550328  
AlaValLysAspIleGluAlaValPro-----ValTrpValAla 212  
Db 1550327 TCACCAAGAGATTGGGACCAATATCCATTACACAGCTGATAAATATGCTATTGGTTGCT 1550268  
ValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLysVal 232  
Db 1550267 TCAACTCTCGGGAGAGATTAAACATA-----TATATGACTATGAGACATTT 1550220  
AlaSer-----TrpIleGluAspLysAspAsnIleLeuLeuTYrGlyThrAspIle-- 249  
Db 1550219 GGAACATCATCTGGAAGA-----ACTGGAAATTTT 1550187  
GluPheIleGlyTYrArg-----AspIleAlaGlyTYr---ArgMetSerValGluGly 266  
Db 1550186 GAGTTTTTGGAGTATTTGGCTTATAGACATAGCTTAAGCATACATATTAGAGTTGTTAAT 1550127  
LeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeuProSerGluLeuLysHis 286  
Db 1550126 GTCAGTAGAGTTGTTGATTAAGACCAAG----- 1550094  
SerGlyArgGluLeuTYrLeuArg-----ThiSerSerTrpAla---ProAspLys 302  
Db 1550093 -----GAGAGGATTTACGTCATGATTTGGCCACGATATCATGGGCTGACACTGAGAGG 1550040  
SerLeuArgIleTrpArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSerTYrAsn 322  
Db 1550039 GATTTAGTCATGGTTA-----GGGAAT---AAGATGCAAGAAATATGTTTGA 1549992  
MetArgGlyGluLeuAla---PheLeuAlaGluAsnSerAspAlaArgGlyTrpGluPro 341  
Db 1549991 AAACCTAAAGATATGTTAATTTATTAAGAAGAACTCAAAACAAATTA----- 1549944  
LeuProGluArgLeuAspAlaPheArgAlaIleTYrAsnAspTrpArg 358  
Db 1549943 -----AAAAAATTAATTAATTTGATGAATATATATAGATGTATTAAG 1549902

```
/ GENERAL INFORMATION:
/ APPLICANT: ASADA, KIYOZO
/ APPLICANT: MUKAI, HIROYUKI
/ APPLICANT: UEMORI, TARASHI
/ APPLICANT: KATO, IKUNOSHIN
/ APPLICANT: LADERMAN, KENNETH
/ APPLICANT: ANFINSEN, CHRISTIAN
/ TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CUSHMAN DABRY & CUSHMAN
/ STREET: 1100 NEW YORK AVENUE, N.W.
/ CITY: WASHINGTON, D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/894,212A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KOKULIS, PAUL N.
/ REGISTRATION NUMBER: 16773
/ REFERENCE/DOCKET NUMBER: 95469/C-1195
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 861-3000
/ TELEFAX: (202) 822-0944
/ TELEX: 6714627 CUSH
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3139 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-07-894-212A-1

Alignment Scores:
Pred. No.: 4.09e-07 Length: 3139
Score: 140.50 Matches: 89
Percent Similarity: 38.11% Conservative: 60
Best Local Similarity: 22.76% Mismatches: 127
Query Match: 7.48% Indels: 115
DB: 1 Gaps: 23

US-09-886-400A-4 (1-364) x US-07-894-212A-1 (1-3139)

QY 25 GtUuYAlAtYrIlleProValIleGtUThLeuIlleYgUgUgUlePro----- 41
DB 800 GAAAGGTGTTACTGCGCGCTTCTTGGAAGACTCTG-----GAGGAATATCCAAACATGAAG 853
QY 42 PheGlyLeuAsnIleThrGlyYrThrLeuYsPheLeuPro-----LysAspIle 58
DB 854 GTTGCCATCATACAGAGTGCGCCCTCATGAGTGCTCCAAATATAGACCCGAATAC 913
QY 59 IleAspLeuValIleGlyLeuIleAlaSerAspLeuIleGluIleIleGlyThrSerTy 78
DB 914 ATAGACTTCTGTTAGAGTCTAGGAAAGAGACAGAGTGAGATGCTGCTGCTGCTTC 973
QY 79 ThrHAlAlleuProLeuProLeuSerArgValGluAlaGlnValGlnArgAsp 98
DB 974 TAGAGCCCTGTGTACATCATATCCCA-----AAGGAATATGAAATAGACAGATA 1024
QY 99 ArgGluValIleGlyLeu-----GluLeuPheGluValSerProGlyYsPheTrpLeuPro 116
DB 1025 AGGTATATGAAAGAGTGCGCTAAGAGTATGATTTGATGCTAGGAGAGTTTGCTAACT 1084
QY 117 GluLeuAlaIleYrAspProIleIleProAlaIleLeuIleYsAsnGlyYrGluYrLeu 136
DB 117 GtUuYAlAtYrAspProIleIleProAlaIleLeuIleYsAsnGlyYrGluYrLeu 136
```

```
DB 1085 GAAAGATATGCGACACAGAGCTCTTAAGACCCCTTAAGAGACCGGAATAGATTATGTA 1144
QY 137 PheAlaAspGlyGluAlaMetLeuPheSerAlaHleuAsnSerAlaIleYsProIle 156
DB 1145 ATAGTTGAC-----GATTACCACTTCATGAGTCCGGGATTAAGTAA 1186
QY 157 LysProLeuTy---ProHleuIleYsAlaGlnArgIle----- 169
DB 1187 GAGAGCTGTACGCGCATATATATACGAAGATGTGGGGAAGTTATAGCTGTTTCCCG 1246
QY 170 -----LysArgPheArgTyIleSerTyLeuLeuGlyLeuArgIleuLeuArgIleYsAla 187
DB 1247 ATAGATGAGAAGTTGAGA-----TATTGATTCCTTTAGACCCGTTGATAGCTC 1297
QY 188 IleYsLeu-----ValPheGluGlyYsValThrLeuYsAlaVal-----Lys 202
DB 1298 TTGAATATACCTGATCTCATAGATGTGATGAGACGAAGATTGCATATTTATCATGAC 1357
QY 203 AspIleGluAlaValProValTrpValAlaValAsnThrAlaValMetLeuGlyIleGly 222
DB 1358 GATGTGAGAAAGTTGGAAATCTGGCCTGGAAGCTTATGAGTGAGTAT----- 1405
QY 223 ArgLeuProLeuMetAsnProLysValAlaSerTrpIle----- 236
DB 1406 -----GAAAGGATGTTTAAAGAAATTTCTTGATAGA 1438
QY 237 -----GluAspIleYsAsnIleLeuLeuTyGlyYrThrAspIleGlu----- 250
DB 1439 ATTTCAAGTATGAAAGATAAAGTAACTTAATGCTTTACAGTAAGTAACTTGAAGAAATATAG 1498
QY 251 -----PheIleGlyYrThrAspIleAlaGlyYrThrArgMetSerValGluYleuLeu 268
DB 1499 CCTAGAGCTTGTGTTATTTCTCAATAGCTTCATAT-----TTT 1537
QY 269 GluValIleAspGluLeuAsnSerGluLeuCysLeuPro----- 281
DB 1538 GAGATG-----AGCAATGCTATTGTCAGCAAGACGCAAGAGCTCTT 1582
QY 282 -----SerGluLeuYsHleuSerGly-----ArgGluLeuTyLeu 293
DB 1583 GTGAGTTGTCGAAGAGCTTAAAGTTAAAGTATATTTGAAAGATACAGGATATTTGTT 1642
QY 294 ArgThrSerSerTrpAlaProAspIleSerLeuArgIleTrpArgGluAspIleGlyAsn 313
DB 1643 AGCGAGGAATTTG-----AAGATTTTCTTATTAATATACCAAGAGACAC 1690
QY 314 -----AlaArgLeuAsnMetLeuSerTyAsnMetArgGlyGluLeuAlaPheLeu 330
DB 1691 TACATGACAAAGAAATGCTATGTAAGTATAGTAAGAAACAT----- 1738
QY 331 AlaGluAsnSerAspAlaArgGlyTrpGluProLeuProGluArgArgLeuAspAlaPhe 350
DB 1739 -----CCTGAGGCCAGGAAGATATCTGCTG 1762
QY 351 ArgAlaIleTyAsnAsp-----TrpArgGly 359
DB 1763 AGAGCAATATGATGATGCTTATTGGACGCG 1795

RESULT 7
US-09-107-532A-1900
/ Sequence 1900, Application US/09107532A
/ Patent No. 6583275
/ GENERAL INFORMATION:
/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
```



```

      ZIP: 02354
      COMPUTER READABLE FORM:
      MEDIUM TYPE: CD-ROM ISO9660
      COMPUTER: PC
      OPERATING SYSTEM: <Unknown>
      SOFTWARE: ASCII
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/107,532A
      FILING DATE: 30-Jun-1998
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 60/085,598
      FILING DATE: 14 May 1998
      APPLICATION NUMBER: 60/051571
      FILING DATE: July 2, 1997
      ATTORNEY/AGENT INFORMATION:
      NAME: Ariniello, Pamela Deneke
      REGISTRATION NUMBER: 40,489
      REFERENCE/DOCKET NUMBER: GTC-012
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (781)893-5007
      TELEFAX: (781)893-8277
      INFORMATION FOR SEQ ID NO: 1900:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 684 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: circular
      MOLECULE TYPE: DNA (genomic)
      HYPOTHEetical: NO
      ANTI-SENSE: NO
      ORIGINAL SOURCE:
      ORGANISM: Enterococcus faecium
      FEATURE:
      NAME/KEY: misc_feature
      LOCATION: (b) LOCATION 1..684
      SEQUENCE DESCRIPTION: SEQ ID NO: 1900:
US-09-107-532A-1900

Alignment Scores:
Pred. No.: 0.00115      Length: 684
Score: 105.00           Matches: 60
Percent Similarity: 37.90%      Conservative: 34
Best Local Similarity: 24.19%    Mismatches: 88
Query Match: 5.59%           Indels: 66
DB: 4                    Gaps: 12

US-09-886-400A-4 (1-364) x US-09-107-532A-1900 (1-684)
Qy 140 G|yG|u|a|a|e|u|p|h|e|s|e|r|a|h|i|s|e|u|a|n|s|e|r|a|h|i|e|l|y|p|r|o|l|e|y|p|r|o|l|e|u 159
Db 7 GGAAGAAAAATGATTTTCAATTAAGAACTGGTG----- 45
Qy 160 T|y|p|r|h|i|s|e|u|l|e|l|y|a|l|a|g|l|n|a|g|l|u|l|y|s|a|t|g|p|h|e|a|t|y|r|i|l|e|s|e|r|y|l|e|u 179
Db 46 -----TTCCGGGTTGATCGTTTATTTATTA 69
Qy 180 G|l|y|e|u|a|g|l|u|e|u|a|g|l|y|s|-----A|a|l|l|e 188
Db 70 GTGGTACGTTGCTGGAGAAATCAGTCAGAAATCACACCGTTGATTTAGTAT 129
Qy 189 l|y|e|u|e|u|a|p|h|e|g|l|u|g|l|y|l|y|s|a|l|t|h|r|e|u|l|y|s|a|l|a|v|a|l|y|s|a|p|----- 203
Db 130 ACGCTGTGTTAGGATCTTAGAAGAAATCAACCTATGATGACAAATGTCATGTAGG 189
Qy 204 -----l|l|e|g|l|a|a|a|l|a|p|r|o|v|a|l|t|r|y|a|l|a|v|a|e|n|t|h|r|a|a|l|a|v|a|l|e|u|-----g|l|y 220
Db 190 CATGTCCTATTTGGCATTGGCGCTTGG-----GCAAGTAATGATTTATGT 234
Qy 221 l|l|e|g|l|a|g|l|e|u|p|r|o|l|e|u|e|a|s|n|p|r|o|l|e|y|s|a|l|a|s|e|r|t|r|i|l|e|g|l|a|s|p|l|y|s|a|p 240
Db 235 ATCGAAGCTATCGTCACAAAATAT---GAAAAGCATCGTTGGGTCAAAAGCGAGGCA 291
Qy 241 A|s|n|l|e|u|e|u|l|y|r|g|l|y|t|h|r|a|s|p|i|l|e|g|l|u|b|h|e|l|l|e|g|l|y|r|a|g|a|s|p|i|l|e|a|g|l|y|r 260

```

```

      : : : : :
      292 TCCTCTCAATCAAAAGAC-----GGCGTCATCAACATGACAGAGTTGACAAATATAC 342
      : : : : :
      261 ArgmetserValG|u|g|l|y|l|e|u|e|u|l|a|l|l|e|a|s|p|i|l|e|u|e|u|s|e|r|g|l|u|e|u|C|y|l|e|u 280
      : : : : :
      343 CATATCGAATGGAGCGATTGCGAGCAATC-----TTAAGACAGCAAGAAATGCTTT 393
      281 Proserglu-----Leu|h|i|s|e|r|g|l|y|a|r|g|l|u|e|u|T|y|r|L|e|u|a|g 294
      394 TTTCTTAAAATGCCAACATGTTATCTTGAGAAATCCCGTCAAAAGAGTGTATGAA 453
      295 Th|s|e|r|s|e|r|t|r|a|p|r|o|a|p|l|y|s|e|r|l|e|u|a|r|g|i|l|e|t|r|p|a|r|g|l|u|a|p|g|i|u|g|l|y|a|s|a|l|a 314
      454 AAATCAGAT-----GAAACAAAGCTTATTCATCTCGCTGTGTATGATGAAGGCAAT 507
      315 Arg|l|e|u|s|n|e|t|l|e|u|s|e|r|T|y|r|a|s|m|e|r|a|r|g|i|l|y|l|u|e|u|a|p|h|e|u|a|l|a|l|u|a|n|s|e|r 334
      : : : : :
      508 CAGCATAAAGCTTCCAGTCCCAATCAATGACTGAA---GCATGCTTATGAGCAACTTG 564
      335 A|s|p|a|l|a|r|g|l|y|t|r|p|l|u|p|r|o|l|e|u|p|r|o|g|l|u|a|r|g|l|e|u|s|p|h|a|p|h|e|a|r|a|l|l|e|T|y 354
      565 AAGAAAGAGGCTAT-----GCAAGTGAATAACAGCTGATTTAT 603
      355 A|s|n|a|s|t|r|p|a|r|g|l|y|l|u|a|e|n|g|l|y 362
      604 GTTGATGTCGACAGAAAAAGG 627

RESULT 8
US-09-103-840A-2
: Sequence 2, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4403765
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 3.31e+03      Length: 4403765
Score: 104.00           Matches: 76
Percent Similarity: 34.82%      Conservative: 49
Best Local Similarity: 21.17%    Mismatches: 128
Query Match: 5.53%           Indels: 106
DB: 3                    Gaps: 18

US-09-886-400A-4 (1-364) x US-09-103-840A-2 (1-4403765)
Qy 61 L|e|u|a|l|y|s|g|l|y|l|e|a|l|a|s|e|r|a|s|p|l|e|u|l|e|g|l|l|e|g|l|t|h|r|s|e|r|T|y|r|H|i|s 80
Db 3386055 CTGCTGGCGCGCGCGATCGACGCCGAGCGTGAAGTGTCTCGTGGCCCACTTGCCAC 3386114
Qy 81 A|l|a|l|e|u|p|r|o|l|e|u|e|u|p|r|o|l|e|u|s|e|r|a|r|g|l|u|a|l|a|g|l|n|a|l|a|g|l|n|a|t|g|a|s|p|a|r|g|l|u 100
Db 3386115 CCGTTCAGCGCGCTGCTG-----GCACCGGCGCTGCGCGAGTTCCGC 3386156
Qy 101 V|a|l|l|e|g|l|u|l|u|e|p|h|e|g|l|u|a|l|s|e|r|-----P|o|l|y|s|g|l|p|h|e|T|p 114
      : : : : :

```

```

Db 3386157 CTGCGGAAAGCCCTCGCGATGCTAGCTCGGCTGCGCGACCCGAAAGGATCTGG 3386216
;
Qy 115 LeuProgluLeuAlaTyrAspProIleIleProAlaIleLeuYsAspAnglyTyrGlu 134
;
Db 3386217 GCACCCGAATGCCATACGCCCGGGAGATGAGTCACTACGCCACCGGGGGTCACT 3386276
;
Qy 135 TyrLeuPheAlaAspGlyGluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIle--- 153
;
Db 3386277 CACTTCATGCTGACAGCGCCCG-----TCGCTGACGCGGACACACCGGCTGGCG 3386324
;
Qy 154 LysProIle----- 156
;
Db 3386325 CGGCGGTGGGGAACCGATGCTGCGCTTCGCTGCGACTTGGAGTCACTACCGG 3386384
;
Qy 157 -----LysProLeuTyrPro---HisLeuIleLysAlaGlnArgGlyLysArg 171
;
Db 3386385 GTGTGTCACCCGAATTCGGCTACCCCGGGCAC-----GCCCTTACCGCGAC----- 3386432
;
Qy 172 PheArgTyrIleSerTyrLeuLeuGlyLeuArgGlyLeuArgGlyLysAlaIleLysLeuVal 191
;
Db 3386433 TTCACACACCTACGACACCTGACCGGACTCAACCGGCGGCTCACCGGCGCTAACGTC 3386492
;
Qy 192 PheGluGlyLysValThrLeuYsAlaValLysAspIleGluAlaValProValTyrVal 211
;
Db 3386493 CCGTCGAGGACA-----AAGGCACTTACGATCCGAGCGCGCTGACCGCGCGTC 3386543
;
Qy 212 AlaValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLys 231
;
Db 3386544 GACGTCACATGTCCTGATTTGTCGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3386594
;
Qy 232 ValAlaSerTyrIleGluAspLysAspAsnIle-----LeuLeuTyrGly 246
;
Db 3386595 -----GAGTCGAGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3386636
;
Qy 247 ThrAspIleGluPheIleGlyTyr----- 254
;
Db 3386637 TTCGACACCGAGTGTTCGCGCACTGTCGTCACGAGGCGCAACTGTCGTCACGAGTA 3386696
;
Qy 255 ---ArgAspIleAlaGlyTyrArgMetSerValGluGlyLeuLeuGluValIleAspGlu 273
;
Db 3386697 TTGCGGAGCTTACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3386756
;
Qy 274 LeuAsnSerGluLeuCysLeuProSerGluLeuLysHisSerGlyArgGlyLeuTyrLeu 293
;
Db 3386757 -----GANTGTCGCGCGACCGCGTCGCAATTG 3386783
;
Qy 294 ArgThrSerSerTyrAlaProAspLysSerLeuArgIleTyrArgGluAspGluYasn 313
;
Db 3386784 CCACCCAGCTCTTGGGGTTCGCGCAAGGACTGCGCAGGTGTGG-----AGCGGTGCC 3386834
;
Qy 314 AlaArgLeuAsnMetLeuSerTyrAsnMetArgGlyGluLeuAla----- 328
;
Db 3386835 AAGGTGGCGATGTGCTCACTCAAC-----AGCGAAGTGTGTCATACCGGTTGACC 3386888
;
Qy 329 -----PheLeuAlaGluAsnSerAspAlaArgGlyTyrGluProLeuProGlu 344
;
Db 3386889 ACCATGCAACAAGCGGTGGCGACAGACGTCCTCGTGAACGA-----CCGCTGCTCGC 3386942
;
Qy 345 ArgArgLeu-----AspAlaPheArgAlaIleTyrAsnAspTyr 357
;
Db 3386943 GATCAGCTTGATGATGATCTCTGCGCAGACCTGCTCACCGTTCACGACGCTGG 3386999
;

```

```

; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

Alignment Scores:
Pred. No.: 3,32e+03 Length: 4411529
Score: 104.00 Matches: 76
Percent Similarity: 34.82 Conservative: 49
Best Local Similarity: 21.17 Mismatches: 128
Query Match: 5.53 Indels: 106
DB: Gaps: 18

US-09-886-400a-4 (1-364) x US-09-103-840A-1 (1-4411529)

Qy 61 LeuValLysGlyLysIleAlaSerAspLeuIleGlyThrSerTyrThrHis 80
;
Db 3390363 CTGTCGCGGCGCTGATCGACCGGACGAGTGAAGTCTGTCGTCGCCACTTGCCAC 3390422
;
Qy 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
;
Db 3390423 CCGTTCACCGCTGCTGTG-----GCACCGCGCGCTGCGCAGTTCGCG 3390464
;
Qy 101 ValLysGluGluLeuPheGluValSer-----ProLysGlyPheTyr 114
;
Db 3390465 CTGCGGAAAGCCTTCGCGCATCTCACTGCGCGCTGCGCGCACCGCGCGAAAGGATCTGG 3390524
;
Qy 115 LeuProGluLeuAlaTyrAspProIleIleProAlaIleLeuYsAspAnglyTyrGlu 134
;
Db 3390525 GCACCCGAATGGCATAGCCCGGGGATGAGGTGCACTACGCCACCGCGGGGTCACT 3390584
;
Qy 135 TyrLeuPheAlaAspGlyGluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIle--- 153
;
Db 3390585 CACTTCATGTCGACGCGCGCG-----TCGTCGACGAGGACACCGCGCTGGGC 3390632
;
Qy 154 LysProIle----- 156
;
Db 3390633 CGGCGGTGGGGAACCGATGTCGCTTCGCTCGCGCATTCGACGTCACGTAACCGG 3390692
;
Qy 157 -----LysProLeuTyrPro---HisLeuIleLysAlaGlnArgGluYasn 171
;
Db 3390693 GTGTGTCACCGAAATCCGCTACCCCGGGCAC-----GCCGCTTACCGCGAC----- 3390740
;
Qy 172 PheArgTyrIleSerTyrLeuLeuGlyLeuArgGlyLeuArgGlyLysAlaIleLysLeuVal 191
;
Db 3390741 TTCACACCTACGACACCACTGACCGGACTCAACCGGCGAGGTGTCACCGGCTAACCGT 3390800
;
Qy 192 PheGluGlyLysValThrLeuYsAlaValLysAspIleGluAlaValProValTyrVal 211
;
Db 3390801 CCGTCGAGCAAC-----AAGGACCTTACGATCCGAGCGGCTGACCGCGCTC 3390851
;
Qy 212 AlaValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLys 231
;
Db 3390852 GACGTCACATGTCGATTTGCTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3390902
;
Qy 232 ValAlaSerTyrIleGluAspLysAspAsnIle-----LeuLeuTyrGly 246
;
Db 3390903 -----GAGTCGAGCGCATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3390944
;
Qy 247 ThrAspIleGluPheIleGlyTyr----- 254
;
Db 3390945 TTCGACACCGAGTGTTCGCGCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3391004
;
Qy 255 ---ArgAspIleAlaGlyTyrArgMetSerValGluGlyLeuLeuGluValIleAspGlu 273
;
Db 3391005 TTGCGGAGCTTACCGCGCGCGGTGTCCGGGTGGCACCTGAGGAGTGCATTCGCCGAC 3391064
;

```

QY 274 LeuanserGluLeuCyseuProserGluLeuLysHisserGlyArgGluLeuTyrleu 293  
 Db 3391065 -----GGATTCGTGCGGAGACCCGGTCAATTG 3391091  
 QY 294 ArgThrSerSerTrpAlaProAspLysSerLeuArgIleTrpArgGluAspGluLysan 313  
 Db 3391092 CCACCACACTCTTGGGTTCCGGCAGACGCTCGCAGAGTGG-----AGCGGTCC 3391142  
 QY 314 AlaArgLeuAsnMetLeuSerTyrAsnMetArgIleGluLeuAla----- 328  
 Db 3391143 AAGTGCCGATCTGGTCTGACCTCAAC-----ACCGAAGTGTGATACCGCGTTGACC 3391196  
 QY 329 -----PheLeuAlaGluAsnSerAspAlaArgIleTyrGluProLeuProGlu 344  
 Db 3391197 ACCATCGACAAAGCGCTGCGCCAGACAGCTCCCTCGACGGA-----CCGCTGCTCGC 3391250  
 QY 345 ArgArgLeu-----AspAlaPheArgAlaIleTyrAsnAspTrp 357  
 Db 3391251 GATCAGCTTGCTGATCAGATCCTGCGGAGACCCCTGCTCACCCTGTCACAGGACTGG 3391307  
 RESULT 10  
 US-08-956-171E-324/c  
 ; Sequence 324, Application US/08956171E  
 ; Patent No. 6593114  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunach  
 ; Gil H. Choi  
 ; Patrick S. Dillon  
 ; Craig A. Rosen  
 ; Steven C. Barash  
 ; Michael R. Fannon  
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 5256  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/956,171E  
 ; FILING DATE: 20-Oct-1997  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/009,861  
 ; FILING DATE: January 5, 1996  
 ; APPLICATION NUMBER: 08/781,986  
 ; FILING DATE: January 3, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mark J. Hyman  
 ; REGISTRATION NUMBER: 46,789  
 ; REFERENCE/DOCKET NUMBER: PB348P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (240) 314-1224  
 ; TELEFAX: (301) 309-8439  
 ; INFORMATION FOR SEQ ID NO: 324:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5030 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 324:  
 US-08-956-171E-324  
 Alignment Scores:  
 Pred. No.: 0.0572 Length: 5030

Score: 103.00 Matches: 58  
 Percent Similarity: 38.6% Conservative: 48  
 Best Local Similarity: 21.17% Mismatches: 90  
 Query Match: 5.48% Indels: 78  
 DB: 4 Gaps: 15  
 US-09-886-400A-4 (1-364) x US-08-956-171E-324 (1-5030)  
 QY 107 GluValSerProLysGlyPheTrpLeuProGluLeuAlaTyrAspProIleIleProAla 126  
 Db 2699 GAGATATCACCT-----TATTATGATTCAATTA----- 2673  
 QY 127 IleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGlyLysAlaMetLeuPheSer 146  
 Db 2672 CTGTAATAATTATATCTACACGCGATATCATTTAAACAGCAGAGAAAAAATGTCACGC 2613  
 QY 147 AlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyrProHisLeuIleLysAla 166  
 Db 2612 TCATTAGAGAAAGCGATATTCGTGCTTAAACTAATTCATTCTTATTAATAGTA 2553  
 QY 167 GluArgGluLysArgPhe-----ArgTyrIleSerTyrLeuLeu 179  
 Db 2552 ATGAAGATATAAAAGTTCAACAGTGTGATTACACAACTAAATTAATGAAGAAACACA 2493  
 QY 180 GlyLeuArgGluLeuArgLysAlaIle-----LysLeuValPheGluGly 194  
 Db 2492 GAACCTTTTCGACATTCAGCCGCTCTAGATAGAGCTAATAACATTAGATATATAGGT 2433  
 QY 195 LysValThrLeuLysAlaValLysAspIleGlu----- 205  
 Db 2432 AATGTAACAATTATATGTTTCCCAAGTTGAGAAAGTCCGAAACGACTATGAAATTA 2373  
 QY 206 ---AlaValProValTrpValAlaValAsnThrAlaValMetLeuGlyIleGlyArg-- 223  
 Db 2372 GCATCAATTCACAACTGATCTTCAAGTAAATCCCTCATTTAGTGTACGAAACAATTG 2313  
 QY 224 LeuProLeuMetAsnProLysLysValAlaSerTrpIleGluAspLysAspAsnIleLeu 243  
 Db 2312 CTGATAGAGTAGGTCCAAAGGCTGAGTGAATGGGTATAAAGACGAGATAGTCTTA 2253  
 QY 244 LeuTyrGlyThrAspIleGluPhe---IleGlyTyrArgAspIleAlaGlyTyrArgMet 262  
 Db 2252 CTA-----ACAGTACACACTTTAGAGATGACACACCAATCATTAATGACACGAGTT 2199  
 QY 263 SerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCyseuProSer 282  
 Db 2198 AGAACTAAGATATGATTAATATGCAATCC----- 2169  
 QY 283 GluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaProAspLys 302  
 Db 2168 -----AAAACACCGGACGTATTTAAAGATGCT 2142  
 QY 303 ---SerLeuArgIleTrpArgGluAspGluLysAsnAlaArgLeuAsnMetLeuSerTyr 321  
 Db 2141 TTCTCATTAAGAAATGTG-----GGCGTCTCATTTGAT--GTGCAATAT 2097  
 QY 322 AsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGluPro 341  
 Db 2096 AAT-----TTCTTAAGAAAGAAACCA-----TGGGAACGA 2067  
 QY 342 LeuProGluLysArgLeuAspAlaPheArgAlaIleTyrAsn 355  
 Db 2066 CTGGAACGTCTAGCT-----AAAGTATTCCAAT 2037  
 RESULT 11  
 US-08-781-986A-324/c  
 ; Sequence 324, Application US/08781986A  
 ; Patent No. 6737248  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunach  
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 5255  
 ; CORRESPONDENCE ADDRESS:

```

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 324:
SEQUENCE CHARACTERISTICS:
LENGTH: 5030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-324

Alignment Scores:
Pred. No.: 0.0572 Length: 5030
Score: 103.00 Matches: 58
Percent Similarity: 38.69% Conservative: 48
Best Local Similarity: 21.17% Mismatches: 90
Query Match: 5.48% Indels: 78
DB: 4 Gaps: 15

US-09-886-400a-4 (1-364) x US-08-781-986A-324 (1-5030)
QY 107 GluValSerProLysGlyPheTrpLeuProGluLeuAlaTyrAspProIleProAla 126
DB 2699 GAGATATCACT-----TATTATGATTCATTA----- 2673
QY 127 IleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGlyGluAlaMetLeuPheSer 146
DB 2672 CTGCTAAATTATCTACACACGCGATATCATTTAAACAAGCAAGAAAAAATGTAACG 2613
QY 147 AlaHisLeuAsnSerLaiIleLysProIleLysProLeuTyrProHisLeuLysAla 166
DB 2612 TCATTACGACAAATCGCTATTCGTGCTTAAACTTAATTCATTCCTTAATTAATGTA 2553
QY 167 GluArgGluLysArgPhe-----ArgTyrIleSerTyrLeuLeu 179
DB 2552 ATGAGAATAATTAAGTTTCAACAAGTGTGATTACACAACATAAATTATTGAAGAAACCA 2493
QY 180 GlyLeuArgGluLeuArgLysAlaIle-----LysLeuValPheGluGly 194
DB 2492 GAACCTTTGACACTTCAGCCGCTCTAGAGTAGAGTACTTAAACATTAGAAATATATAGGT 2433
QY 195 LysValThrLeuLysAlaValLysAspIleGlu----- 205
DB 2432 AATGTAACAATTATATGCTTTCCAAATGTGAGAAAAGTCGCAACCAAGATATGAAATTA 2373
QY 206 ---AlaValProValTrpValAlaValAsnThrAlaValMetLeuGlyTyrArg--- 223
DB 2372 GCATCAATTCCAACGTATCTTCAAGTAATAATGCTTCATTATAGTGTACGAAACAATTG 2313
QY 224 LeuProLeuMetCAsnProLysLysValAlaSerTrpIleGluAspLysAspAsnIleLeu 243
```

```

DB 2312 CTTGATGAAGTAGTGCCTCAAAAGGTGTAGCTGAATGGGTTAAAAAGCAGATGATGCTTA 2253
QY 244 LeuTyrGlyThrAspIleGluPhe---IleGlyTyrArgAspIleAlaGlyTyrArgMet 262
DB 2252 CTA-----ACAGATCAACACTTAAAGATGACACACCAATCAATTATTAAGTACACAGATT 2199
QY 263 SerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeuProSer 282
DB 2198 AGAACTAAGAGATATGATTAATATCGCATCC----- 2169
QY 283 GluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaProAspLys 302
DB 2168 -----AAACACGCGACGCTATTTAAAGATGCT 2142
QY 303 ---SerLeuArgIleTrpArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSerTyr 321
DB 2141 TTCTCACTGAAGATGTGG-----GGCGGTGTACTATTCAT---GTGCGATAT 2097
QY 322 AsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTrpGluPro 341
DB 2096 AAT-----TTCTTGAAGAAACCA-----TGGGAACGA 2067
QY 342 LeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsn 355
DB 2066 CTGGAACGTCTACGT-----AAAGCTATTCCAAT 2037

RESULT 12
US-08-270-013B-1
Sequence 1, Application US/08270013B
Patent No. 5686294
GENERAL INFORMATION:
APPLICANT: Sogabe et al.
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
NUMBER OF INVENTION: 2
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 61601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/270,013B
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 164701/1993
FILING DATE: 02-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Robert F.
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 62321
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: (25)3533
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus stearothermophilus
```

STRAIN: ATCC12016  
US-08-270-013B-1

## Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.0133	102.50	37.54%	21.84%	5.46%	1912	64	46	92	92	14

US-09-886-400a-4 (1-364) x US-08-270-013B-1 (1-1912)

```

QY 10 LeuGlnTyrAlaGluIleProLysSerGluIleProLys----- 22
Db 1070 ATGACGTAACGCTTTAAGAAATCCGATTCGCAAAAACCGGTCATCGCCACATCG 1129
QY 23 ---ValIleGluYsaIaTyrIleProValIleGluThrLeuIleGluGluIlePro 41
Db 1130 GCGGCTTGATACGCGCGC-----TTCCGACGTTGCTGCGCGAGAG----- 1174
QY 42 PheGlyLeuAenIleThrGlyTyrThrLeuYsPheLeuProLysAspIleIleAspLeu 61
Db 1175 -----CTGAACATTTG-----GTAAAGATGTCATCGGCTTT 1207
QY 62 ValIysGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThiAla 81
Db 1208 GTTTAGGCGGCGCATGGGATGACATGCTGCGCTGCTCGTACTGTCGCGCGCGC 1267
QY 82 Ile-----LeuProLeuLeuProLeuSerArgValGluAlaGluValIleAspArg 99
Db 1268 ATTCGCGTCGAAAAATCATTCCGAAAGATCGTTTGAAGCGCATGTTGACGCGAGCGC 1327
QY 100 GluValIysGluGluLeuPheGluValSerProLysGly-----PheTyrLeuProGlu 117
Db 1328 AAAGCGCGCGGTAATGTCACCTGCTCGCAACGCGACGCGCTACTACGACCGCGC 1387
QY 118 LeuAlaTyrAspProIleIleProAlaIleLeuYsAspAn----- 131
Db 1388 GCCTGCTTGTGCAAAATGTCGAAACGATTGTAAGAACAGCGCGCATTTTGGCGCG 1447
QY 132 -----GlyTyrGluTyrLeuPheAlaAspGlyGluAlaMet 143
Db 1448 ATGCGCTACCTTGAAGCGGATACGCGCTATGAGGCAATTATTGGCGTCCGACGATC 1507
QY 144 LeuPheSerAlaHISLeuAsnSerAlaIleYs----- 154
Db 1508 CTCGCGCGCAACGCGCATCGAAGATGATCGAGCTCGAGCTGACCGAAGAGAAAAACG 1567
QY 155 -----ProIleYs----- 157
Db 1568 CGCTGCGCAAAATCCGTCGAATCCGTTAAAAATGTCATCGCATCGTGAATAGCGCGAG 1627
QY 158 -----ProLeuTyrProHISLeuIleYsaIaGlnArgGluYsArgPhe----- 172
Db 1628 GCAAAATATCCGCGATTCGCGGATTTTGTCCCACTCATATGAAGCGCTTTCTAGAC 1687
QY 173 ---ArgTyrIleSerTyrLeuLeuGlyLeuArgGluLeuArgYsaIaIle----- 188
Db 1688 AACGAGGCGGTGGAGAC-ATGTTGAAAAACGAAGCTCGGAGACGCGATCGGGAAATT 1746
QY 189 -----LysLeuValPheGluGlyIlyValThrLeuYsaIaValIlyAspIle 204
Db 1747 CAAGCGGCGGAAAAAGCTGTTCCAAAGCCGCAATC-----GAAGCAAAAGACTTG 1797
QY 205 -----GluAlaValProValTyrValAlaValAsnThrPala 216
Db 1798 CTTCTTATATCTTGAGCTGACGAGATGATCCCAATCGGCTATATCCAGCATGATTATGCT 1857
QY 217 ValMetLeuGlyIleGlyArgLeuProLeuMetAsnPro 229
Db 1858 TCACAGACGCGCTTTGAAA-----CGGCTGTCGCGCGC 1893

```

## RESULT 13

US-08-838-418-1

Sequence 1, Application US/08838418

Patent No. 574342

GENERAL INFORMATION:

APPLICANT: Sogabe et al.

TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE

TITLE OF INVENTION: DEHYDROGENASE ACTIVITY

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Leydig, Voit & Mayer, Ltd.

STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601-6780

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/838,418

FILING DATE: 17-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/270,013

FILING DATE: 01-JUL-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 164701/1993

FILING DATE: 02-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hoover, Allen E.

REGISTRATION NUMBER: 37354

REFERENCE/DOCKET NUMBER: 78339

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5600

TELEFAX: (312) 616-5700

TELEX: (25) 3533

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1912 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Bacillus stearothermophilus

STRAIN: ATCC12016

US-08-838-418-1

## Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.0133	102.50	37.54%	21.84%	5.46%	1912	64	46	92	92	14

US-09-886-400a-4 (1-364) x US-08-838-418-1 (1-1912)

```

QY 10 LeuGlnTyrAlaGluIleProLysSerGluIleProLys----- 22
Db 1070 ATGACGTAACGCTTTAAGAAATCCGATTCGCAAAAACCGGTCATCGCCACATCG 1129
QY 23 ---ValIleGluYsaIaTyrIleProValIleGluThrLeuIleGluGluIlePro 41
Db 1130 GCGGCTTGATACGCGCGC-----TTCCGACGTTGCTGCGCGAGAG----- 1174
QY 42 PheGlyLeuAenIleThrGlyTyrThrLeuYsPheLeuProLysAspIleIleAspLeu 61

```

```
Db      1175 -----CTGAACATTTCG-----CTAAAGATGTCACGTGGGTTT 1207
QY      62  VallyGlyIyllealaserpleuilegullelythrserlyrthrhisala 81
Db      1208  GTTTTAGGCGCCATGCGCATGTCATGTCGCCCTCGCTCGCTACCTCGCGCGCGC 1267
QY      82  Ile-----LeuProleuLeuProleuSerArgValGluIalaglnValGlnArgAspArg 99
Db      1268  ATTCGCGCTCGAAAACTCATCTCCGAAAAGATCGTTTGAGCCCATCGTTGAGCGGAGCGC 1327
QY      100  GluVallyleGluIleuPheGluValSerProlyleGly-----PheTyrleuProGlu 117
Db      1328  AAGCGCGCGCGTGAATCGTCAACCTGCTCGGACAGCGGAGCGCTACTACGACCGCGCC 1387
QY      118  LeuAlaTyrAspProlelleleProAlalleuIylAspAsn----- 131
Db      1388  GCCTCGCTGTGCAAAATGTCGAAGCATTTTGAAGACCGCGCGCATTTTGC CGCGC 1447
QY      132  -----GlyTyrGluTyrleuPheAlaAspGlyIalMet 143
Db      1448  ATGCGCTACTTGAAGCGCAATACGGCTATGAAGGCATTTATTTGGCGCTCGCGAGCATC 1507
QY      144  LeuPheSerAlaHisleuAsnSerAlalleys----- 154
Db      1508  CTCGCGCGGCAACGGCATCGAAGAAAGTGCAGCTGAGCTGACCGAAGAGGAAAAAGCG 1567
QY      155  -----ProIleIys----- 157
Db      1568  CGCTCGCAAAATCCGTCGAATCCGTTAAATAATGTCATCGCATCGTGAATAGCGCGAG 1627
QY      158  -----ProleuTyrProHisleuIleValaglnArgGluIylAspArgpe----- 172
Db      1628  GCAAAATTCGCGCATTTGCCGATTTTCTCCACAGTCAATGAAGACCTTTCTTAGAC 1687
QY      173  --ArgTyrIleSerTyrleuLeuGlyLeuArgGluIylAspAlaIle----- 188
Db      1688  AACGAAGGGGTGGGAAC-ATGTTGAAAAAACGAAGCTCGGAGACCGATCGGGGAAT 1746
QY      189  -----LysleuValPheGluGlyLysValThrleuIylAspAlaIylAspIle 204
Db      1747  CAAGCGGGGAAAAAGTCGTGTTCCAAAGCCGCATC-----GAAGAAGAAAGCTTG 1797
QY      205  -----GluAlaValProValTyrValAlaValAsnThrAla 216
Db      1798  CTTCTTATCTTGCGGTGAGCGATGATGCCATCCGCTATATCCAGCATGATATGCT 1857
QY      217  ValMetleuGlyIleGlyArgLeuProleuMetAsnPro 229
Db      1858  TCACAGACGCCGTTTGGAAAA---CCGCTCGTCCCGCG 1893

RESULT 14
US-09-710-279-2149
; Sequence 2149, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: F0348005
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2149
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2149
```

```
Alignment Scores:
Pred. No.: 0.00693 Length: 1083
Score: 101.50 Matches: 57
Percent Similarity: 36.45% Conservative: 52
Best Local Similarity: 19.06% Mismatches: 117
Query Match: 5.40% Indels: 73
DB: 4 Gaps: 12

US-09-886-400a-4 (1-364) x US-09-710-279-2149 (1-1083)
QY      4  LeuValPheHisGlyAsnleuGlnTyrAlaGluIleProlySerGluIleProlyVal 23
Db      181  CTTATTTCAAGGA-----CTAAAGCATATC 207
QY      24  IleGluValAlaTyrIleProValIleGluThrleuIleGluGluIleProPheGly 43
Db      208  GTTGAAGATGATATGATGATTAATTCGTAACCTTAATCCAAATATATATC-----AAT 261
QY      44  LeuAsnIleThrGlyTyrThrleuIylSerProlyAspIleIleAspLeuVallys 63
Db      262  CTATATGCAATTACACATATCTTGATGTAATCTTAAGGTGTCATGCAATGTTAGCG 321
QY      64  GlyGlyIleAlaserAspleuileGluIleIleGlyThr-----SerTyrThrhisala 81
Db      322  GATCAAAATGCTTTAGAGAACATATCATGATTAATACAAATAGCTCATATTTATTAACA 381
QY      82  IleuProleuLeuProleuSerArgValGluAlaGlnValGlnArgAspArgGluVal 101
Db      382  GTTCAAACTTTATATACCTTAATAATTAATTAAGAA-----GATTTTC 420
QY      102  LysGluGluLeuPheGluValSerProlyGlyPheTyrleuProGluIleuAlaTyrAsp 121
Db      421  AAGACAGTTTAAACGAACCTT----- 441
QY      122  ProIleIleProAlaIleleuIylAspAsnGlyTyrGluTyrleuPheAlaAspGlyGlu 141
Db      442  -----GGATTAACCTTAAGAAAGTAATTAAGATATGTTCTTTGAAAGTGA 489
QY      142  AlaMetleuPheSerAlaHisleuAsnSerAlalleysProIleIylSerProleuTyrPro 161
Db      490  GGT-----AAAGGCAATTTTAAACGATGATGATGCAAGTCCCT 528
QY      162  HisleuIleValleuIalaglnArgLysArgPheArgTyrIleSerTyrleuGlyLeu 181
Db      529  TATATGGGAAGTTAGATAGT-----ATCGAATATGTT----- 561
QY      182  ArgGluLeuArgLysAlalleysleuValPhe-----GluGlyLysVal 196
Db      562  -----GATGAATTAATAAATTGAGTTTATGATTAAGACAAATGAAATTAAGAAAT 609
QY      197  ThrleuValAlaValIylAspIleGlu-----AlaValProVal-----TyrValAla 212
Db      610  ACTAAACGTGCTATTTAGATATATACCCCATACGAAACACAGTTTGTATTTATTAATA 669
QY      213  ValAsnThrAlaValMetleuGlyIleGlyArgLeuProleuMetAsnProlyIylVal 232
Db      670  ATGAACAAAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
QY      233  AlaSerTyrIleGluAspLysAspAsnIleleuLeuIylSerGlyThrAspIleGluPheIle 252
Db      730  TTAGATGAATTTTCTGAATATGCAAAAAACAGCTCAATATACCGAGCGTACGATATACA 789
QY      253  GlyTyrArgAspIleAlaGlyTyrArgMetSerValGluGlyLeuGluVal----- 270
Db      790  GGTCAACATGATATGCAATTAAGAAAGTATGATGATGATGATGATGATGATGATGATGAT 849
QY      271  -----IleAspGluLeuAsnSerGluLeuCysleuProSerGluLeuIylHis 286
Db      850  GAGTATAAAGCTTAGCCAACTTGAGCAGATGTTTGTATGCTGATATTAATTAACAC 906

RESULT 15
US-09-710-279-2105
```

Sequence 2105, Application US/09710279  
 Patent No. 6703492  
 GENERAL INFORMATION:  
 APPLICANT: KIMBERLY, WILLIAM JOHN  
 TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 FILE REFERENCE: PUS48005  
 CURRENT APPLICATION NUMBER: US/09/710,279  
 CURRENT FILING DATE: 2000-11-09  
 PRIOR APPLICATION NUMBER: 60/164,258  
 PRIOR FILING DATE: 1999-11-09  
 NUMBER OF SEQ ID NOS: 4472  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 2105  
 LENGTH: 1131  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 US-09-710-279-2105

Alignment Scores:  
 Pred. No.: 0.00745 Length: 1131  
 Score: 101.50 Matches: 57  
 Percent Similarity: 36.45% Conservative: 52  
 Best Local Similarity: 19.06% Mismatches: 117  
 Query Match: 5.40% Indels: 73  
 DB: 4 Gaps: 12

US-09-886-400a-4 (1-364) x US-09-710-279-2105 (1-1131)

```

Qy      4 LeuValPheHisGlyAsnLeuGlnTrpAlaGluIleProLysSerGluIleProLysVal 23
Db      139 CTATTTTCAAGGA-----GTAACGATAC 165

Qy      24 IleGluValAlaTrpIleProValIleGluTrpLeuIleLysGluIleIleProPheGly 43
Db      166 GTTGAGAGTGAATGATGATATATGCTAACTTATCCAAATAATATC-----AAT 219

Qy      44 LeuAsnIleThrGlyTrpThrLeuLysPheLeuProLysAspIleIleAspLeuValLys 63
Db      220 CTATAGCATTAACACATTAATCTGATGTAATCTTAAGGTCATCGAATGTTACCG 279

Qy      64 GlyGlyIleAlaSerAspLeuIleGluIleIleGlyThr-----SerTrpThrHisAla 81
Db      280 GATCAATAAGGTTTAAAGAACATGATCAATGATTAATACCAATAGCTCATATTATTACAAA 339

Qy      82 IleLeuProLeuLeuPheLysSerArgValGluAlaGlnValGlnArgAspArgGluVal 101
Db      340 GTTCAAACTTTTATATCTTAAATAATTATTTGAA-----GATTTC 378

Qy      102 LysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTrpAsp 121
Db      379 AAAGCAGTTTAAACGAACTT----- 399

Qy      122 ProIleIleProAlaIleLeuLysAspAsnGlyTrpGluTrpLeuPheAlaAspGlyGlu 141
Db      400 -----GGATTAGCTAAAGAAAGTAATTAAGATATGTTCTTGAAAGTGA 447

Qy      142 AlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTrpPro 161
Db      448 GGT-----AAAGGCAATTTAAACCGATGAGTGAATGCAATGCTT 486

Qy      162 HisLeuIleLysAlaGlnArgGluLysArgPheArgTrpIleSerTrpLeuLeuGlyLeu 181
Db      487 TATATAGGGAAGTATGATAGT-----ATCGAATATGTT----- 519

Qy      182 ArgGluLeuArgLysAlaIleLysLeuValPhe-----GluGlyLysVal 196
Db      520 -----GATGAATATAAACTTGATTAATGATTAAGCAATGAATTAAGAAATA 567

Qy      197 ThrLeuLysAlaValLysAspIleGlu-----AlaValProVal-----TrpValAla 212
  
```

```

Db      568 ACTAAACGTGCTATTTTATGATAATCACCACATACGAAACACCGATTGATTTTATTAATA 627
Qy      213 ValAsnThrAlaValMetLeuGlyTrpIleArgLeuProLysMetAspProLysVal 232
Db      628 ATGAACAAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Qy      233 AlaSerTrpIleGluAspLysAspAsnIleLeuLeuTrpGlyTrpAspIleGluPheIle 252
Db      688 TTAAGTGAATTTTCTGAATATGCCCCAAAAACAGCTCAATATACCGAGGTACGATATACA 747
Qy      253 GlyTrpArgAspIleAlaGlyTrpArgMetSerValGluGlyLeuLeuGluVal----- 270
Db      748 GGTCAACATGATAGTCCATTAAGAAAGTACCTATCATAGGTGTTGAGTATGATGATTT 807
Qy      271 -----IleAspGluLeuAsnSerGluLeuCyLeuProSerGluLeuLysHis 286
Db      808 GAGTATAAAGCTACCAACTTGAGAGCAGATGTTTGTACTGATGATTAACAC 864
  
```

Search completed: January 28, 2005, 06:04:35  
 Job time : 3883 secs

**This Page Blank (uspto)**